

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MParch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 10 07:03:26 1996; MasPar time 11.71 Seconds

607.114 Million cell updates/sec

Tabular output not generated.

Title: >US-08-446-915-2  
Description: (1-409) from US08446915.pep

Sequence: 1 MASSAPDENEFQGCPPAP.....KHAYVKDDTMIKCIVD TSA 409

Scoring table: PAM 150

Gap 11

Searched: 49340 seqs, 17385503 residues

Post-processing: Minimum Match 0%

### Listing first 45 summaries

Database: swiss-prot32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9

Statistics: Mean 49.616; Variance 101.312; scale 0.490

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2945	100.0	409	8	TRF1_MOUSE	TNF RECEPTOR ASSOCIAT	0.00e+00
2	972	33.0	501	8	TRF2_MOUSE	TNF RECEPTOR ASSOCIAT	3.50e-16
3	148	5.0	760	5	MEPA_MOUSE	MEPRAIN A ALPHA-SUBUNIT	7.90e-07
4	127	4.1	1048	7	SBCC_ECOLI	EXONUCLEASE SBCC.	7.06e-04
5	120	4.1	2291	7	SPCB_DROME	SPECTRIN BETA CHAIN.	6.00e-03
6	115	3.9	1319	3	DY15_DROME	150 KD DYNEIN--ASSOCIA	2.65e-02
7	112	3.8	493	4	INVO_SARGO	INVOLICRIN.	6.33e-02
8	112	3.8	870	6	POL_TSRV	POLYPROTEIN (REVE	6.32e-02
9	113	3.8	1427	7	REST_HUMAN	RESTIN (CYTOPLASMIC L	4.78e-02
10	109	3.7	544	4	INVO_MOUSE	INVOLICRIN.	1.49e-01
11	106	3.6	668	5	MEPB_RAT	MEPRAIN A BETA-SUBUNIT	3.44e-01
12	105	3.6	776	1	ARCB_ECOLI	AEROBIC RESPIRATION C	4.55e-01
13	107	3.6	2653	2	ACEB_HUMAN	CENTROMERIC PROTEIN E	2.61e-01

15	102	3.5	428	4	INVO_CEBL	INVOIDOLIN.	1.02e+00
14	103	3.5	443	6	PC17_HUMAN	PARAMOPLASTIC CEREBE	7.82e-01
16	103	3.5	467	3	FUMC_ECOLI	FORMATE HYDRATASE CL	7.82e-01
17	104	3.5	741	1	BSG2_DROME	BLASTOGEN SPECIFIC P	5.96e-01
18	103	3.5	872	8	ULF4_HVSD	97 KD ALPHA TRANS-IND	7.82e-01
19	103	3.5	2017	5	MYSN_DROME	MYOSIN HEAVY CHAIN, N	7.82e-01
20	100	3.4	92	9	Y108_SSVI	POLYHEDRAL 10.8 KD	1.75e+00
21	99	3.4	312	9	VPHE_HSPVI	POLYHEDRAL ENVELOPE P	2.27e+00
22	99	3.4	333	8	VG24_NSVI	HYPOTHETICAL GENE 24	2.27e+00
23	101	3.4	522	4	INVO_HYLLA	INVOIDOLIN.	1.34e+00
24	100	3.4	560	4	INVO_PANPA	INVOIDOLIN.	1.75e+00
25	101	3.4	567	7	RECK_ECOLI	DNA REPAIR PROTEIN RE	1.34e+00
26	100	3.4	567	5	NRAP_HAEIN	SENSOR PROTEIN NARO H	1.75e+00
27	101	3.4	917	7	SLAP_THETH	S-LAYER PROTEIN PRECU	1.34e+00
28	99	3.4	1053	3	DYNA_CHICK	DYNACTIN, 117 KD ISOE	2.27e+00
29	101	3.4	1335	3	CYNA_RAT	DYNACTIN, 150 KD ISOE	1.34e+00
30	101	3.4	1505	2	CCAT_HUMAN	CCAT DISPLACEMENT PR	1.34e+00
31	99	3.4	2833	4	IP3B_DROME	INOSITOL 1,4,5-TRISP	2.27e+00
32	98	3.3	236	8	VC05_SPVVA	HYPOTHETICAL PROTEIN	2.95e+00
33	98	3.3	312	1	CA4H_HUMAN	CARBONIC ANHYDRASE IV	2.95e+00
34	98	3.3	363	1	BASS_ECOLI	SENSOR PROTEIN BASS/P	2.95e+00
35	98	3.3	450	4	IFEB_HELPO	NON-NEURONAL CYTOPL	2.95e+00
36	98	3.3	458	8	VIM1_XENLA	VIMENTIN 1 AND 2.	2.95e+00
37	98	3.3	564	5	M12_STRPY	M PROTEIN, SEROTYPE 1	2.95e+00
38	96	3.3	575	4	IFEB_HELAS	NON-NEURONAL CYTOPL	4.95e+00
39	97	3.3	577	5	MEP1_NECOR	MITOCHONDRIAL PROCESS	3.83e+00
40	98	3.3	876	5	MYSN_HUMAN	MYOSIN HEAVY CHAIN, S	2.95e+00
41	98	3.3	962	2	COPB_DROME	COTOMER BETA SUBUNIT	2.95e+00
42	98	3.3	974	5	MYSB_MESAU	MYOSIN HEAVY CHAIN, C	3.83e+00
43	97	3.3	1589	6	PHP_DROME	POLYHEDROTIC-PROXIMAL	3.83e+00
44	97	3.3	1938	5	MYSN_AQRI	MYOSIN HEAVY CHAIN, S	4.95e+00
45	96	3.3	1959	5	MYSN_CHICK	MYOSIN HEAVY CHAIN, N	4.95e+00

## ALIGNMENTS

ID	RESULT	1	STANDARD;	PRT;	409 AA.
AC	TRE1	MOUSE			
AC	P39428;				
DT	01-FEB-1995	(REL. 31, CREATED)			
DT	01-FEB-1995	(REL. 31, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)			
DE	TNF RECEPTOR ASSOCIATED FACTOR 1 (TRAF1).				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
LN	(1)				
RN	SEQUENCE FROM N A., AND SEQUENCE OF 123-135 AND 390-402.				
RX	MEDLINE; 9434937L.				
RA	ROTHE M., WONG S.C., HENZEL W.J., GOEDEL D.V.;				
RL	CELL 75:681-692(1994).				
CC	- - FUNCTION: SIGNAL TRANSDUCER ASSOCIATED WITH THE CYTOPLASMIC DOMAIN				
CC	OF THE 75 KD TUMOR NECROSIS FACTOR RECEPTOR (TNF-R2).				
CC	- - SUBCELLULAR LOCATION: CYTOPLASMIC.				
CC	- - SUBUNIT: HETERODIMER OF TRAF1 AND TRAF2.				
DR	EMBL; L35302; L35302.				
FM	COILED COIL.				
SQ	SEQUENCE 409 AA; 4546 MW; 857196 CN;				
Query Match	100.0%;	Score 2945;	DB 8;	Length 409;	
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;			
Matches 409;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	







DE WOLF-PETERS C. SHIPMAN B.

CC -!- FUNCTION: INVOLUCRIN IS A KERATINOCYTE PROTEIN THAT FIRST APPEAR

CC IN THE CELL CYTOSOL, BUT ULTIMATELY BECOMES CROSS-LINKED TO





50 SEQUENCE 443 AA; 50549 MW; 1032055 CN;

Matches 30; Conservative 30; Mismatches 44; Indels 11; Gaps 10;

Db 13 lqgdlqlaelgklldrntele-dsvqmytngelqeielyltkvel lrgmnehak 71

Qy 134 LERNIS-ELQI-QAAVEATGILEVDCIRAPCESQEEL-ALQHLVKE-KULAQI-EEKIR 188

72 vyeql-dvtareleetnqklvadskasq-qk--ilsltetieclqtnidhlgsgv 122

Qy 189 VFANIVAVINKEVEASHALAAISHQSLDREHLISLEQRVEIQQTAAQ-KDQV 242

Job time : 16 secs.

Job time : 16 secs.



(7M)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

```

March_gp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Dec 10 07:03:59 1996; MasPar time 17.22 Seconds
Tabular output not generated. 603,847 Million cell updates/sec

```

Title:	>US-08-446-915-2
Description:	(1-409) from US08446915.pep
Perfect Score:	2945
Sequence:	1 MASSAPDENEFGCCPAP .....KHAYKDDTFLKCIYDTSA 409

Scoring table: PAM 150

Searched: 82130 seqs, 25426960 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

```
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unann9 13:unperc
14:unrev
```

Statistics: Mean 47.486; Variance 124.260; scale 0.382

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being compared, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2589	87.9	416	11	B55649	TNFR-associated prot	0.00e+00	
2	956	32.5	501	12	S56163	tumor necrosis facto	3.02e-12	
3	794	27.0	568	11	A55649	TNFR-associated prot	9.73e-10	
4	148	5.0	760	12	A40195	meprin A (EC 3.4.24.	6.60e-05	
5	144	4.9	748	12	S24134	endopeptidase 2 (EC	1.93e-04	
6	127	4.3	1048	3	BVE5CS	sbc protein - Esche	1.57e-02	
7	120	4.1	2291	5	A46147	spectrin beta chain	8.87e-02	
8	119	4.0	714	11	S39464	pABA peptide hydrola	1.13e-01	
9	115	3.9	1319	10	A28313	glut protein - fruit	2.96e-01	
10	115	3.9	75239	11	S37539	macroglobin - human	2.96e-01	
11	115	3.9	3259	11	A56539	giantin - human	2.96e-01	
12	111	3.8	700	11	S49383	meprin A (EC 3.4.24.	7.59e-01	

Dec 10 06:47

US-08-446-915-2.mpr

45	13	112	3.8	870	3	GNAVY	pol polyprotein - eh	6.01e-01
44	14	113	3.8	1427	11	S2295	restin - human	4.75e-01
43	15	110	3.7	613	10	S27170	hypothetical protein	9.58e-01
42	16	106	3.6	601	11	S33377	P63 protein - human	2.40e+00
41	17	106	3.6	668	4	A42908	meprin A (EC 3.4.24,	2.40e+00
40	18	107	3.6	704	12	A48040	meprin beta chain pr	1.91e+00
39	19	105	3.6	778	3	R3CEAR	aerobic respiration	3.00e+00
38	20	105	3.6	1392	11	A43336	microtubule-vesicle	3.00e+00
37	21	107	3.6	2663	2	S28261	kinesin-related prot	1.91e+00
36	22	103	3.5	342	5	A24263	myosin heavy chain,	4.70e+00
35	23	104	3.5	385	5	C24263	myosin heavy chain,	3.76e+00
34	24	103	3.5	443	11	A49833	autocalcineion recogn	4.70e+00
33	25	103	3.5	467	1	UFEC	fumarate hydratase (	4.70e+00
32	26	103	3.5	509	11	A40448	DNA-binding protein	4.70e+00
31	27	104	3.5	582	10	S24545	intermediate filament	3.76e+00
30	28	104	3.5	622	12	A57281	kinesin-like motor p	3.76e+00
29	29	104	3.5	622	12	S52687	ChO2 antigen - Chine	3.76e+00
28	30	104	3.5	741	10	A26572	bsg5D protein - fru	3.76e+00
27	31	103	3.5	872	3	TNBEH	97K alpha trans-indu	4.70e+00
26	32	102	3.5	1992	5	A47297	myosin heavy chain f	5.86e+00
25	33	103	3.5	2017	2	A36014	myosin heavy chain,	4.70e+00
24	34	100	3.4	92	7	S03214	hypothetical protein	9.09e+00
23	35	100	3.4	169	8	S03744	hypothetical protein	9.09e+00
22	36	101	3.4	567	3	R0EEN	reep protein - Esche	7.31e+00
21	37	100	3.4	567	8	F64058	nitrate sensor prote	9.09e+00
20	38	101	3.4	738	9	A48246	ethylene-response pr	7.31e+00
19	39	101	3.4	917	7	S26565	P100 protein - Therm	7.31e+00
18	40	99	3.4	1053	11	A41642	dynamacin - chicken	1.13e+01
17	41	100	3.4	1093	11	A47212	TATR element modul	9.09e+00
16	42	100	3.4	1151	5	S09331	myosin beta heavy ch	9.09e+00
15	43	101	3.4	1325	10	A51629	dynein-associated pr	7.31e+00
14	44	101	3.4	189	10	A45592	liver stage antigen	7.31e+00
13	45	99	3.4	2833	10	A43360	inostol 1,4,5-tiltip	1.13e+01

## ALIGNMENTS

RESULT	1
ENTRY	B55649 #type complete
TITLE	TNFR-associated protein EB16 - human
ORGANISM	#formal name Homo sapiens #common name man
DATE	23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change

ACCESSIONS  
REFERENCE

#authors  
Mosialos, G.; Birkenbach, M.; Yalamanchilli, R.; Vahrsdale  
T.; Ware, C.; Kieff, E.

## Journal

**Abstract** The Epstein-Barr virus transforming protein LMP1 engages signaling proteins for the tumor necrosis factor receptor family.

## # accession

```
##status      preliminary
##molecule_type  mRNA
##residues      1-416 ##label MOS
##cross-references  GB:U19261
```

## SUMMARY

Query Match 87.9%; Score 2589; DB 11; Length 416;

Matches 345; Conservative 38; Mismatches 25; Indels 1; Gaps 1;

[illegible]

Qy	2	ASSANDENEEFGCPPARCQBPSEPRVLCITACLSLNADDEDRICPRORADNHLHPSP	61
Db	68	gsrlrteqekahpveaagiqcpfagvgcsfkgspsvgehvetsqstshnlhlgfmmqwk	127
Qy	62	CSPL-TQEKVSHDVAEAEIMCFAGVGCSPKSPSQMOEHEATSSGSHLLAVAIKEMK	120
Db	128	arlgqcleagpmaleqnlbedlqqaavearqdlvdcytrapceesqeelqhfimkell	187
Qy	121	SFSGNULSGAMALERNLSQLQAAVEATGDLVEDOCYRAPCGESQEEALMQUVEXKL	180
Db	188	aesegklsfenviavlnkveaahalatsishsqdlreilslsegrvvelsgtlaqkd	247
Qy	181	AQLEEKARFANIVAVLNKEVASHLAAASTHQSJDREHLLSTGEORVVELQQTIAQKD	240
Db	248	qalqllegqlhmeasfdgtflwkintvrrrrchesacqtvslisapalytakygkcl	307
Qy	241	QVLAKLEHSLTMEASFDGTFIMKINVTKRCHESYCGRTVSLSPAFYTAKEYKCLT	300
Db	308	rlylmgdqgktrchlsflvimrmgeydaallpwpfinkvtfmllqdmrnehaidafpdis	367
Qy	301	RYLYMDSGKKTLSLFLVIMRGETYALLPWPFRNKVTFMLLQDNREHAIIDAFRPDIS	360
Db	368	saeafqrpqseecvnaagcpdlffpleklsqphayvddtmfllkcolvetat	416
Qy	361	SAFQRPOSSETNVAASGCPLEFFLSRLQDSRHAAYKVDJTMETLACIVDSTA	409

ENTRY	2	RESULT
TITLE	556163	#type complete
ALTERNATE_NAMES	tumor necrosis factor receptor-associated protein - human	
ORGANISM	TNF receptor-associated protein	
DATE	#formal_name Homo sapiens #common_name man	
	10-Oct-1995 #sequence_revision 01-Dec-1995	#text_change
ACCESSIONS	01-Dec-1995	
REFERENCE	556163; S58925; S58926	
	556163	
#author	Song, H. Y.; Donner, D. B.	
#journal	Biochem. J. (1995) 309:825-829	
#title	Association of a RING finger protein with the cytoplasmic domain of the human type-2 tumour necrosis factor receptor	
#accession	S56163	
#molecule_type	mRNA	
#residues	1-501 ##label SON	
#cross-references	EMBL:U12597	
REFERENCE	S58925	
#author	Song, H. Y.; Donner, D.	
#description	submitted to the EMBL Data Library, July 1994	
	Association of a RING finger protein with the cytoplasmic domain of the human type 2 TNF receptor.	
#accession	S58925	
#molecule_type	mRNA	
#residues	1-42,63-342,363-501 ##label SON	
#cross-references	EMBL:U12597	
#accession	S58926	
#molecule_type	mRNA	
#residues	1-342,'REPOQCCHRVCSFCLASIL',363-501 ##label SOF	
#cross-references	EMBL:U12597	
KEYWORDS	zinc finger	
SUMMARY	2	
	#length 501	
	#molecular-weight 55842	#checksum 5395

Query Match	32.5%	Score 956;	DB 12;	Length 501;
Best Local Similarity	53.6%;	Pred. No. 3.02e-125;		
Matches 120;	Conservative	58;	Mismatches 45;	Indels 1;
			Gaps 1;	

Db 275 lëktatfenivcwlInnervrantæacacqrhlldqdkleasakvqdetsiglkda 334  
 Qy 183 lEEkLIRfANIVAVlNKEVEASHLALASIHQSOLDREHLSlEQRVELlQDTLAKQKQV 242  
 Db 335 madleqvlemeasdydvfyfiksldprklæavagrldafispafergykmcrlr 394  
 Qy 243 lGKLEHSlALMEASFDGFIWKrlTNYTKRCHESVCGRVYSlEPAFTTKYCKICRL 302  
 Db 395 ylngdgqrqthslæffvymkgpndallwprfnqrvlmlldgmrehvldafpds 454  
 Qy 303 YlNDGSGGKTHlSLFVlMRGEYDlALLPAPPRNKVTYFMLLDQNNREHlADAFRDLSSA 362  
 Db 455 sfqrpyndmiasgcpfæpovekneæ-knyvrdallkalyd 497  
 Qy 363 sfQRQSEtNVASGCPfEPRlSKLDSPRAHYVQDYMfLKCIVD 406

RESULT	3	
ENTRY	A55649	#type complete
TITLE	TNFR-associated protein LMP1 - human	
ALTERNATE_NAMES	CD40-binding protein	
ORGANISM	#formal_name Homo sapiens	#common_name man
DATE	23-Mar-1995	#sequence_revision 23-Mar-1995
	05-Apr-1995	#text_change
ACCESSIONS	A55649; A55135	
REFERENCE	A55649	
	Moslatai, G.; Birkenbach, M.; Yalamanchili, R.; VanArsdale, F.; Ware, C.; Kieff, E.	
	Cell (1995) 80:389-399	
#journal	The Epstein-Barr virus transforming protein LMP1 engages	
#title	signaling proteins for the tumor necrosis factor receptor	
	family.	
#accession	A55649	
#molecule_type	mRNA	
#residues	1-568	#label MOS
#cross-references	GB:U015650	
#note	nucleotide sequence not given	
REFERENCE	A55135	
#authors	Hu, H.M.; O'Rourke, K.; Boguski, M.S.; Dixit, V.M.	
#journal	J. Biol. Chem. (1994) 269:30069-30072	
#title	A novel RING finger protein interacts with the cytoplasmic	
	domain of CD40.	
#accession	A55135	
#molecule_type	mRNA	
#residues	1-133, 135-404, 'C', 406-568	#label HUA
#cross-references	GB:U015637	
#note	nucleotide sequence not given	
KEYWORDS	coiled coil	
FEATURE	#region RING-finger motif	
53-91	#length 568 #molecular-weight 64490 #checksum 8660	
SUMMARY		

Query Match	27.0%	Score 794	DB 11	Length 568
Best Local Similarity	38.8%	Pred. No. 9,73e-100		
Matches 149	Conservative	93	Mismatches 118	Indels 24
			Gaps 18	
<hr/>				
Db	189	lqhtedcdp-cvwsophkcaqvctllrselashlsecmatpccsfkrycvqgtnq	247	
	:	:	:	:
	:	:	:	:
	:	:	:	:
QY	40	lRDEDRICPCGRADN-lHPVSPGSPILTOE-KVH-SDVAELATMCFAGVGCSCFKSPGS	96	
	:	:	:	:
	:	:	:	:
	:	:	:	:
Db	248	lkaheasavqhvml-----kewsnlekvsl  qmesveknksisq hndi-csfefsi	302	
	:	:	:	:
	:	:	:	:
	:	:	:	:
QY	97	MOEHAFOSSHLVLLAVIKEMKSPSPNSGADPM-ALENNISELQLOAAVETDQLEY	155	
	:	:	:	:
	:	:	:	:
	:	:	:	:







KEYWORDS	endonuclease; hydrolase; nucleotidyltransferase; polypeptide; reverse transcriptase
SUMMARY	#length 870 #molecular-weight 99312 #checksum 4310
Query Match	3.8%; Score 112; DB 3; Length 870;
Best Local Similarity	24.7%; Pred. No.6.0e-01;
Matches 24; Conservative	27; Mismatches 42; Indels 4; Gaps 4;
Db 187	qlylvhymdilllahtcdhl-lyga-f-silkghslnglviadekigtthfpynglgsly 244
Qy 166	ELMAGHUKKELIAGLEEKIAVFANIVANUKKEVASHALAAASHQSUDREH-LSLE 226
Db 245	prvynqtqvlkqtch-lkrlntdfkl1gdimwirpyl 280
Qy 227	QRVEIQQTLAQKQOVLGKLEHSIRIMEASFDGTF 263
RESULT 14	SZ2695 #type complete
ENTRY TITLE	restin - human
ORGANISM	#formal_name Homo sapiens #common_name man
DATE	04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Jun-1993
ACCESSIONS	SZ2695; S19853
REFERENCE	SZ2695
#authors	Blibe, G.; Delabie, J.; Bruengen, J.; Richener, H.; Aasselberg, F.A.M.; Cerletti, N.; Sorg, C.; Odink, K.; Tarceay, L.; Wiesendanger, W.; DeWolf-Peetere, C.; Shipman, R.
#journal	EMBO J. (1992) 11:2103-2113
#title	Restin: a novel intermediate filament-associated protein highly expressed in the Reed-Sternberg cells of Hodgkin's disease.
#cross-references	MOID:92289675
#accession	SZ2695
#molecule	type mRNA
#residues	1-1427 ##label BIL
#cross-references	EMBL:X64838
SUMMARY	#length 1427 #molecular-weight 160989 #checksum 428
Query Match	3.8%; Score 113; DB 11; Length 1427;
Best Local Similarity	25.0%; Pred. No.4.75e-01;
Matches 24; Conservative	31; Mismatches 34; Indels 7; Gaps 7;
Db 451	tkgdlctqtkleharl-kelqas-l-fkctkadtqgledtrrvatvseksimelekd 508
Qy 165	SOEELIAQ-HLVKEKILQALEEKIAVFANIVAV-LNKVEVASHLA-LAASHQSUDREH 221
Db 509	alrvge-vaelttrrl-esmkpavgdmslellgeis 542
Qy 222	LSLEQRVEIQQTLAQKQOVLGKLEHSIRIMEAS 257
RESULT 15	SZ7770 #type fragment
ENTRY TITLE	hypothetical protein 1 - African malaria mosquito (fragment)
ORGANISM	#formal_name Anopheles gambiae #common_name African malaria mosquito
DATE	17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Jun-1993
ACCESSIONS	SZ7770
REFERENCE	SZ7770
#authors	Bessansky, N.J.; Paskewitz, S.M.; Mills-Hamm, D.M.; Collins,

[illegible]

**Accession** F07700

**Submission** submitted to the EMBL Data Library, June 1992

**Description** Distinct families of site-specific retroposons occupy identical positions in the RNA genes of *Anopheles gambiae*

```
##molecule_type DNA
##residues 1-613 ##label BCS
##cross-references EMBL:M9360
#length 613 #checksum 108
```

Query Match	3.7%;	Score 110;	DB 10;	length 613;
Best Local Similarity	32.0%;	Pred. No. 9.58e-01;		
Matches	24;	Conservative	17;	Mismatches 31;
			Indels	3;
			Gaps	3;

[illegible]

Search completed: Tue Dec 10 07:04:22 1996  
Job time : 23 secs.

Dec 10 06:51

US-08-446-915-4.pir

1

US-08-446-915-4

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MSPrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 10 07:07:14 1996; MafPar time 20.07 Seconds

634.805 Million cell updates/sec

Tabular output not generated.

Title: >US-08-446-915-4  
Description: (1-501) from US08446915.pep  
Perfect Score: 3702  
Sequence: 1 MAASVTSFGSELLQPEFS.....NSVWDMAIFIKALVDLTGL 501

Scoring table:

PAM 150

Gap 11

Searched: 82130 seqs, 25426960 residues

Post-Processing: Minimum Match 0%

Listing first 45 summaries

Database: pir47

1:am1 2:am2 3:am3 4:unann1 5:unann2 6:unann3 7:unann4  
8:unann5 9:unann6 10:unann7 11:unann8 12:unann9 13:unannc  
14:unrev

Statistics: Mean 48.057; Variance 117.898; scale 0.408

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the total score being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	3770	88.3	501	12	S56163 tumor necrosis facto	0.00e+00
2	986	26.6	416	11	B55649 TNFR-associated prot	1.74e-138
3	806	21.8	568	11	A55649 TNFR-associated prot	3.03e-108
4	229	6.2	458	9	A29361 DC17 protein - sline	3.43e-16
5	162	4.4	700	11	S49383 meprin A (EC 3.4.24.	4.44e-07
6	162	4.4	760	12	A40195 meprin A (EC 3.4.24.	4.44e-07
7	158	4.3	487	3	DDBY18 DNA repair protein R	1.44e-06
8	158	4.3	704	12	A48040 meprin beta chain pr	1.44e-06
9	157	4.2	668	4	A42908 meprin A (EC 3.4.24.	1.93e-06
10	155	4.2	748	12	S24134 endopeptidase 2 (EC	3.45e-06
11	153	4.1	1483	3	RGBYH1 regulatory protein H	6.15e-06
12	153	4.1	1502	14	S59400 CYC1/CYP3 transcript	6.15e-06

Dec 10 06:51

US-08-446-915-4.pir

2

13	141	3.8	317	11	A47380 RING finger-containi	1.86e-04
14	139	3.8	714	11	S39464 PABA peptide hydrola	3.24e-04
15	135	3.6	658	9	A44272 photomorphogenesis r	9.74e-04
16	126	3.4	336	11	S06573 finger protein (clon	1.10e-02
17	127	3.4	355	10	S57398 NIVA protein - Emeri	8.40e-03
18	127	3.4	443	10	S57328 uvw8 protein - Emeri	8.40e-03
19	127	3.4	477	10	S55494 C3Hc4 type zinc fing	8.40e-03
20	120	3.2	160	10	S28290 hypothetical protein	5.24e-02
21	119	3.2	227	5	S05585 tropomyosin - human	6.77e-02
22	119	3.2	596	10	S51157 intermediate filament	6.77e-02
23	120	3.2	1863	11	A54652 breast/ovarian cance	5.24e-02
24	114	3.1	284	11	S23470 beta-tropomyosin - A	2.40e-01
25	114	3.1	385	5	C24263 myosin heavy chain,	2.40e-01
26	113	3.1	400	7	S28022 outer membrane prote	3.08e-01
27	114	3.1	501	14	S34825 uvw2 protein - Neuro	2.40e-01
28	110	3.0	529	11	S06565 finger protein (clon	6.45e-01
29	110	3.0	675	11	S51037 zinc-finger protein	6.45e-01
30	112	3.0	1042	12	S42511 RAG-1 protein - rabb	3.94e-01
31	112	3.0	1042	12	S44379 RAG-1 protein - rabb	3.94e-01
32	111	3.0	1043	12	B33754 recombination-activa	5.05e-01
33	110	3.0	1043	11	A33754 recombination-activa	6.45e-01
34	111	3.0	1992	5	A47297 myosin heavy chain f	5.05e-01
35	107	2.9	321	11	D39371 Ig V-region-like B-g	1.33e+00
36	109	2.9	393	12	JN0533 finger protein pML2-	8.23e-01
37	107	2.9	406	14	S59296 probable Zn-finger (	1.33e+00
38	106	2.9	458	5	S09228 intermediate filament	1.65e+00
39	107	2.9	879	5	S20486 paramyosin - fruit f	1.33e+00
40	107	2.9	879	5	S20208 paramyosin, standard	1.33e+00
41	104	2.8	284	2	TWCHS1 tropomyosin 1, smoot	2.77e+00
42	104	2.8	458	5	A43549 vimentin 1 - African	2.72e+00
43	105	2.8	463	5	B43549 vimentin 4 - African	2.15e+00
44	104	2.8	624	11	S28418 probable zinc-bindin	2.72e+00
45	104	2.8	2094	11	S33124 tpr protein - human	2.72e+00

#### ALIGNMENTS

RESULT	1	
ENTRY	S56163	#type complete
TITLE	tumor necrosis factor receptor-associated protein - human	
ALTERNATE_NAMES	TNF receptor-associated protein	
ORGANISM	#formal name Homo sapiens #common name man	
DATE	10-Oct-1995 #sequence #revision 01-Dec-1995 #text_change 01-Dec-1995	
ACCESSIONS	S56163; S58925; S58926	
REFERENCE	S56163	
##authors	Song, H.Y.; Donner, D.B.	
##journal	Biochem. J. (1995) 309:725-873	
##title	Association of a RING finger protein with the cytoplasmic domain of the human type-2 tumour necrosis factor receptor.	
##accession	S56163	
##molecule_type	mRNA	
##residues	1-501 ##label SON	
##cross-references	EMBL:012597	
REFERENCE	S58925	
##authors	Song, H.Y.; Donner, D.	
##submission	submitted to the EMBL Data Library, July 1994	
##description	Association of a RING finger protein with the cytoplasmic domain of the human type 2 TNF receptor.	
##accession	S58925	
##molecule_type	mRNA	
##residues	1-42,63-342,363-501 ##label SON	
##cross-references	EMBL:012597	
##accession	S58926	



Dec 10 06:51

US-08-446915-4.pdf



```
##molecule_type mRNA
##residues 1-342,'RPFDQCGHRYCSFCLASIT',363-501 ##label SOF
##cross-references EMBL:U12597
KEYWORDS
zinc finger
SUMMARY
#length 501 #molecular-weight 55842 #checksum 5395
```

	Query Match	88.3%;	Score 3270;	DB 12;	Length 501;	
	Best Local Similarity	86.6%;	Pred. No. 0.00e+00;			
	Matches	434;	Conservative	36;	Mismatches 31;	Indels 0; Gaps 0;
D <sub>b</sub>	1 maaasvtpggslellpqgfsktllgtklieakyicsacenvlrrpfagcdghryscflas	60				
O <sub>y</sub>	1 MAASVTPSGSLELPDQGFSKTLGTKLIEAKYICACAKNLRPFAGCDGHRYSCFLTS	60				
D <sub>b</sub>	1 llsagqmcacayhegiiveegisllssasafpdnaarrevsiipavcsdcctwktllke	120				
O <sub>y</sub>	61 ILSAGQMCACAYHEGIIVEEGISILSSSAFPDNAREVESIIPAVCSDCCTWKTKLKE	120				
D <sub>b</sub>	121 yeschgcgcmpltecpackgyivrltgkethelhepcorslschrzpcpgcadvtahhe	180				
O <sub>y</sub>	121 YESCHGCGLPFLTCECPACKGLIVRLTSEKHTHEDPCPKRSISQHCRAPCSNDLEVAYE	180				
D <sub>b</sub>	181 vcskfntcgcggfkkipretfqdnvttcgkcrvprthaigalevevegkgdhevyql	240				
O <sub>y</sub>	181 VCSKFNTCGCGFKKIPRETFQDNVTTGKCRVPRTHAIGALEVEVEKGDEHYQL	240				

RESULT	2	
ENTRY	B55649	#type complete
TITLE	TNR-associated protein EB16 - human	
ORGANISM	#formal name Homo sapiens #common name man	
DATE	23-Mar-1995	#sequence_revision 23-Mar-1995 #text_change 23-Mar-1995
ACCESSIONS	B55649	
REFERENCE	A55649	
#authors	Mosialoe, G.; Birkenbach, M.; Yalamanchili, R.; Vanharsdale, T.; Ware, C.; Kleff, E.	
#journal	Cell (1995) 80:389-399	
#title	The Epstein-Barr virus transforming protein LMP1 engages signaling proteins for the tumor necrosis factor receptor family.	
#accession	B55649	
#status	preliminary	
#molecule_type	mRNA	

Dec 10 06:51

US-08-446-915-4.rpt



```
##residues      1-416 ##label MOS
##cross-references GB:U19261
SUMMARY      length 416  #molecular-weight 46163  #checksum 68153
```

Query Match	26.6%	Score 986	DB 11	Length 416
Best Local Similarity	55.4%	Pred. No. 1.74e-138		
Matches 124	Conservative 59	Mismatches 40	Indels 1	Gaps 1
Db	190	leqglvfenlavlvekeeahalatslshgsqlderslslegrvvelqgltaqkdda	249	
		:                : :    :       :    :    :		
Qy	275	LEQGIATFENIVCYLHNEVERAAVTAACSCGRHRLDQPKLEALSNKVOQLERSIGLKDDA	334	
Db	250	lqklegslmeasfagqfllklntvrrchessacqvtvslfspaftyakgylclrl	309	
		::    : :    :    :    :    :    :    :    :    :    :    :    :    :    :    :		
Qy	335	MADLEQKVELEVSTIDGCFIKLSIDPTRKQDAVAGCTTAITSFPAFTSYRGTKMKLV	394	
Db	310	ylngdqgrtrchlsflfvmrgvdtallpwfnkvrtfmlldmnehaidafpdlasa	369	
		:       :    :		
Qy	395	YLNQDGTGRCHLSLFVWKKQPDALLQPPENQKVTLMALDNNHNEHVDAERPDVTS	454	
Db	370	sfgqpgsetevnasqblfplqlkdsphkqvvdctmflklciue	413	
		:    :          :    :    :    :    :    :    :    :    :    :    :    :		
Qy	455	SFGPVSQDNILASGPIECGPKSMEA-KASIVYDAITITAIYVD	497	

RESULT	5	
ENTRY	S49383	#type complete

REFERENCE  
#authors A41196  
Dummermuth, E., Serchi, E.E., Jiang, W., Wolz, R.L., Bond,  
J.S., Flannery, A.V., Beynon, R.J.  
#journal J. Biol. Chem. (1991) 266:21381-21385  
#title The astacin family of metalloendopeptidases.  
#cross-references M01D:9204:028

Dec 10 06:51

US-08-446-915-4.rpt

7

```
#accession B41196
##molecule_type mRNA
##residues 77-275 ##label DUM
##cross-references GB:M74897; GB:M74238
CLASSIFICATION #superfamily astacin homology; MAM homology
FEATURE
1-33 #domain signal sequence #status predicted #label SIG\
34-77 #domain propeptide #status predicted #label PRO\
78-760 #product meprin A alpha chain #status experimental
#label MAT\
87-273 #domain astacin homology #label AST\
276-445 #domain MAM homology #label MAM
SUMMARY #length 760 #molecular-weight 85702 #checksum 4733

Query Match 4.4%; Score 162; DB 12; Length 760;
Best Local Similarity 32.9%; Pred. No. 4,44e-07;
Matches 28; Conservative 25; Mismatches 28; Indels 4; Gaps 4;

Db 447 vwtlnslqglentvkqdk-lv-sprfyne-gygyvrylpmgrltsnsglqltthly 503
:| | : : : : : : : : | | | | : : : : : : : : : : : : : : : : : :
Qy 335 IMKISDFRKQEVAVGRTPAIFSPAYTSRYGKMGCLRVYINAGDGTGRGTHLSIFVVM 414
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 504 sgnddallqwpncvqaintllqde 528
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 415 KGPNDALLQMPFNQKVTLM-LLDHN 438

RESULT 7
ENTRY DDBY18 #type complete
TITLE DNA repair protein RAD18 - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein YCR066w
ORGANISM #formal name Saccharomyces cerevisiae
DATE 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change
12-May-1995

ACCESSIONS S05802; S22263; S19481; JS0082
REFERENCE S05802
#authors Jones, J.S.; Weber, S.; Prkash, L.
#journal Nucleic Acids Res. (1988) 16:7119-7131
#title The Saccharomyces cerevisiae RAD18 gene encodes a protein
that contains potential zinc finger domains for nucleic
acid binding and a putative nucleotide binding sequence.
#cross-references MUID:88303333
#accession S05802
##molecule_type DNA
##residues 1-487 ##label JON
##cross-references EMBL:X12588
REFERENCE S22260
#authors Benit, P.; Chanet, R.; Fabre, F.; Faye, G.; Fukuhara, H.;
Sor, F.
#journal Yeast (1992) 8:147-153
#title Sequence of the sup61-RAD18 region on chromosome III of
Saccharomyces cerevisiae.
#cross-references MUID:92221691
#accession S22263
##molecule_type DNA
##residues 1-487 ##label BEN
##cross-references EMBL:S93798
REFERENCE S19477
#authors Antoine, G.; Benit, P.; Chanet, R.; Fabre, R.; Faye, G.;
Fukuhara, H.; Mathieu, A.; Sor, F.
#submission submitted to the Protein Sequence Database, March 1992
#accession S19481
#molecule_type DNA
#residues 1-487 ##label ANT
```

Dec 10 06:51

US-08-446-915-4.rpt

8

```
#cross-references EMBL:X59720
REFERENCE JS0082
#authors Chanet, R.; Magana-Schwencke, N.; Fabre, F.
#journal Gene (1988) 74:543-547
#title Potential DNA-binding domains in the RAD18 gene product of
Saccharomyces cerevisiae.
#cross-references MUID:8923745
#accession JS0082
##molecule_type DNA
##residues 1-487 ##label CHA
GENETICS
#gene LISA:RAD18
#map_position 3R
CLASSIFICATION #superfamily DNA repair protein RAD18
KEYWORDS ATP; DNA binding; DNA repair; zinc finger
FEATURE
28-48 #region zinc finger motif\
51-65 #region zinc finger motif\
190-210 #region zinc finger motif\
366 #binding site ATP (lys) #status predicted
SUMMARY #length 487 #molecular-weight 55230 #checksum 1135

Query Match 4.3%; Score 158; DB 3; Length 487;
Best Local Similarity 36.7%; Pred. No. 1,44e-06;
Matches 18; Conservative 16; Mismatches 14; Indels 1; Gaps 1;

Db 21 qlctllrchickdflkvpvltpcgtfscalcitlhnp-npclele 68
:| : | | : : : : : : | | : : : : : : : : : : : : : : : : : :
Qy 27 RLEAKTLCGACNIIIRPQACGARYCSFCLTILSSGPNACATYE 75

RESULT 8
ENTRY A48040 #type complete
TITLE meprin beta chain precursor - mouse
ORGANISM #formal name Mus musculus #common name house mouse
DATE 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change
07-Jul-1995

ACCESSIONS A48040
REFERENCE A48040
#authors Gorbea, C.M.; Marchand, P.; Jiang, W.; Copeland, N.G.;
Gilbert, D.J.; Jenkins, N.A.; Bond, J.S.
#journal J. Biol. Chem. (1993) 268:21035-21043
#title Cloning, expression, and chromosomal localization of the
mouse meprin beta subunit.
#accession A48040
##status preliminary
##molecule_type mRNA
##residues 1-704 ##label GOR
##cross-references GB:I15193
CLASSIFICATION #superfamily MAM homology; astacin homology
FEATURE
72-258 #domain astacin homology #label AST\
261-430 #domain MAM homology #label MAM
SUMMARY #length 704 #molecular-weight 79585 #checksum 7013

Query Match 4.3%; Score 158; DB 12; Length 704;
Best Local Similarity 32.7%; Pred. No. 1,44e-06;
Matches 34; Conservative 34; Mismatches 22; Indels 14; Gaps 11;

Db 432 lwhlqgfc---q-llgqgtsvypfyysk-gy-a-fqlym-dlr-ystnnglyfhl 481
| | | : | | : : : : : : : : | | : : : : : : : : : : : : : : : :
Qy 355 IMKISDFRKQEVAVGRTPAIFSPAYTSRYGKMGCLRVYINAGDGTGRGTHLSIFVVM 414

Db 482 sgnddqlqwpncvqaintllqmpdlrqmfng-rlsttdpt 524
```

	I::           : ::  : :  :    :  : :  :
Oy	415 KGPNDALQMPFN-QKVTMLLDHN-N-REHVIADFRDVTSSS 455
RESULT	9
ENTRY	A42908 #type complete
TITLE	mepirin A (EC 3.4.24.18) beta chain - rat
ALTERNATE_NAMES	endopeptidase-2 beta chain; mepirin beta chain; mepirin-a beta chain
ORGANISM	#formal_name Rattus norvegicus #common name Norway rat
DATE	17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Mar-1995
ACCESSIONS	A42908
REFERENCE	A42908 Johnson, G.D.; Hersch, L.B.
#authors	J. Biol. Chem. (1992) 267:13505-13512
#journal	Cloning a rat mepirin cDNA reveals the enzyme is a heterodimer.
#title	
#cross-references MUID:92317075	
#accession	A42908
#status	preliminary
#molecule_type	mRNA
#residues	1-668 ##label JOH
#cross-references NCBI:P:107784	
#note	sequence extracted from NCBI backbone
CLASSIFICATION	#superfamily mepirin A; actacin homology; MAM homology
KEYWORDS	heterodimer; hydrolase
FEATURE	
72-258	
261-430	#domain actacin homology #label ASTV
SUMMARY	#domain MAM homology #label MAM
	#length 668 #molecular-weight 75049 #checksum 9806
Query Match	4.2%; Score 157; DB 4; Length 668;
Best Local Similarity	40.0%; Pred. No. 1.93e-06;
Matches	34; Conservative 23; Mismatches 16; Indels 12; Gaps 9,
Db	432 iwhipft---g--llggct--tysaplysak-gy-a-fqnl--dltp-tnyghl 480    :      : ::  : :  :    :  : :  :
Oy	335 IMKISDFRKDEAVNRPALFSAFYSTRGYCMCLRYLIMGDTGRTHLSIFVVM 414
Db	481 sganddqlpwcpwgatnllidqn 505  ::          : ::  : :  :    :  : :  :
Oy	415 KGPNDALQMPFN-QKVTMLLDHN 438
RESULT	10
ENTRY	S24134 #type complete
TITLE	endopeptidase 2 (EC 3.4.24.-) - rat
ALTERNATE_NAMES	endopeptidase 24.18
ORGANISM	#formal_name Rattus norvegicus #common name Norway rat
DATE	02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 19-Oct-1995
ACCESSIONS	S24134
REFERENCE	S24134 Corbelli, D.; Gaudoux, F.; Mainwright, S.; Ingram, J.; Kenny, A.J.; Boileau, G.; Cline, P.
#authors	FEBS Lett. (1992) 309:203-208
#title	Molecular cloning of the alpha-subunit of rat endopeptidase-24.18 (endopeptidase-2) and co-localization with endopeptidase-24.11 in rat kidney by in situ hybridization.
#cross-references MUID:9231675	
#accession	S24134
#status	preliminary

	#molecule_type mRNA	
CLASSIFICATION	#residues 1-748 ##label COR	
KEYWORDS	#superfamily MAM homology; astacin homology hydrolase; metalloprotein; proteinase; zinc FEATUDE	
FEATURES	76-262 #domain astacin homology ##label ASTV 265-434 #domain MAM homology ##label MAM 156,160,166 #binding site zinc (His) #status predicted	
SUMMARY	#length 748 #molecular-weight 85138 #checksum 2333	
Query Match	4.2%; Score 159; DB 12; Length 748;	
Best Local Similarity 31.8%; Pred.No.3.45e-06;		
Matches 27; Conservative 23; Mismatches 31; Indels 4; Gaps 4;		
Db	436 wtlinisqvlehtvky-drl-y-sprifnse-gyfgvlypngritsmsyglqlafily 492	
QY	355 IWKISDFTKKQGVAGRPDAFSPATYSRYGRKMCLRVYLINGUGRGCHLSFFVM 414	
Db	493 sgndhvilewpeuegintltldge 517	
QY	415 KGPNDALIQWEPFNOKVTLM-LIDHN 438	
RESULT 11		
ENTRY	RGBYH1 #type complete	
TITLE	regulatory protein HAP1 - yeast ( <i>Saccharomyces cerevisiae</i> )	
ALTERNATE_NAMES	regulatory protein CYP1	
ORGANISM	#formal name <i>Saccharomyces cerevisiae</i>	
DATE	30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 13-Sep-1995	
ACCESSIONS	AJ31312; S15447; S05804; S15446	
REFERENCE	AJ31312	
#authors	Pfeiffer, K.; Kim, K.S.; Kogan, S.; Guarente, L.	
#journal	Cell (1989) 56:291-301	
#title	Functional dissection and sequence of yeast HAP1 activator.	
#cross-references	MUID:89106221	
#accession	AJ31312	
#molecule_type DNA		
#residues	1-1483 ##label PFE	
#cross-references	EMBL:J03152	
REFERENCE	S15447	
#authors	Creusot, F.; Verdriere, J.; Gaisne, M.; Slonimski, P.P.	
#journal	J. Mol. Biol. (1988) 204:263-276	
#title	CYP1 (HAP1) regulator of oxygen-dependent gene expression in yeast. I. Overall organization of the protein sequence displays several novel structural domains.	
#cross-references	MUID:89125585	
#accession	S15447	
#molecule_type DNA		
#residues	1-62, 'R', 64-144, 'T', 146-322, 'R', 324-454, 'S', 456-507, 'V', 509-586, 'N', 588-882, 'D', 884-959, 'H', 961-1150, 'D', 1152-1156, 'S', 1158-1304, 'N', 1305-1483 ##label CREI	
#cross-references	EMBL:X13793	
#note	the sequence is from mutant CYP1-18	
GENETICS		
#gene	LISTA:CYP1; HAP1	
CLASSIFICATION	#superfamily regulatory protein HAP1; GAL4 zinc binuclear cluster homology	
KEYWORDS	DNA binding; heme binding; zinc finger	
FEATURES	#domain DNA-binding #status predicted #label DNA\ #domain GAL4 zinc binuclear cluster homology #label GAL4\	
	#region zinc finger CCCG modify	
64-84		

Dec 10 06:51

US-08-446915-4.pdf

**1997**

	177-189	#region glutamine-rich\
	245-445	#domain heme-binding #status predicted #label HEM\
	299-304, 323-328,	
	347-352, 373-378.	
	389-394, 415-420	#region 6-residue repeat\
	1308-1403	#domain activation element #status predicted #label ACTN
	1388-1483	#region acidic
SUMMARY	#length 1483 #molecular-weight 164270 #checksum 1441	
Query Match	4.1%; Score 153; DB 3; Length 1483;	
Best Local Similarity	31.6%; Pred. No. 6,15e-06;	
Matches	25; Conservative 23; Mismatches 29; Indels 2; Gaps 1;	
D <sub>B</sub>	59 rlpdsctcrirkvkcdklrphcgqcktgvaahlymedtwaaeeakelkldnekklr i18 :: ::     :: : ::    :     :: :             :	
O <sub>y</sub>	184 KFLPTDGGCGKKKIPRETRQDHVRAAGSKRV--LCRHVTYGCSEAVETENVDHLEQLRRL 241	
D <sub>B</sub>	119 ervksjeltlekvhyspsse i37  ::   :: : :::	
O <sub>y</sub>	242 EHIALLLSFTLEAASPGT 260	

ENTRY	12
TITLE	S59400 #type complete CYC1/CYP3 transcription activator - yeast (Saccharomyces cerevisiae)
ORGANISM	#formal name Saccharomyces cerevisiae
DATE	30-Nov-1995 #sequence_revision 30-Nov-1995 #text_change

ACCESSIONS	559400
REFERENCE	559386

Accession	Authors	Submission	Description
F01492	Johnson, D.	submitted to the EMBL Data Library, February 1995	The sequence of <i>S. cerevisiae</i> cosmid 9672.

```

#status      preliminary
#residues    1-1502  ##label JOH
#cross-references  EMBL:U07865
#length 1502  #molecular-weight 166106  #checksum 76698
SUMMARY

```

Query Match	4.1%;	Score 153;	DB 14;	Length 1502;
Best Local Similarity	31.6%;	Pred. No. 6.15e-06;		
Matches	25;	Conservative	23;	Mismatches 29;
			Indels	2;
			Gaps	1;

[illegible]

RESULT	13
ENTRY	A47380
TITLE	#type complete
ORGANISM	RING finger-containing DNA binding protein RING1 - human
DATE	#format name Homo sapiens #common name man
	21-Jan-1994
	#sequence_revision 18-Nov-1994
	#text_change 18-Nov-1994

ACCESSIONS	REFERENCE
A47380	Lowering, R.; Hanson, I.M.; Borden, K.L.; Martin, S.; O'Reilly, N.J.; Evan, G.I.; Rahman, D.; Pappin, D.J.; Trowsdale, J.; Freemont, P.S.
A47380	

Dec 10 06:51

US-08-446-915-4.rpt

12

```

#journal      Proc. Natl. Acad. Sci. U.S.A. (1993) 90:2112-2116
#title        Identification and preliminary characterization of a protein
               motif related to the zinc finger.
#cross-references MUID:93211912
#accession    A47380
#status        preliminary
#molecule_type DNA; protein
#residues      1-377 ##label LOV
#cross-references NCBI:128010; NCBI:128011
#note          sequence extracted from NCBI backbone
#length 377 #molecular-weight 39145 #checksum 7895
SUMMARY

```

Query Match	3.68	Score 141	Db 11	Length 377
Best local Similarity	3.85	Pred. No. 1,866-04		
Matches	15	Conservative	14	Mismatches 16; Indels 1; Gaps 1
D	13	lheslmpcislmlkmtkthcrlfrfscslvtrlsrgnkeepc	58	
Y	28	lslarylcsacsknlrrpfqa-dcgrrfscslslssgpncaac	72	

RESULT ENTRY	14
TITLE	S39464 #type complete
ORGANISM	PABA peptide hydrolase alpha chain - human
DATE	#format_name Homo sapiens #common_name man 19-May-1994 #sequence_revision 17-Nov-1995 #text_change 01-Dec-1995
ACCESSIONS	S39464
REFERENCE	S39464

#authors	Dumerm
	Ster

#Journal  
#title  
FEBS Lett. (1993) 335:367-375  
Cloning of the PABA peptide hydrolase alpha subunit  
(PPH-alpha) from human small intestine and its expression  
in COS-1 cells.

```
##accession      S39464
##status         preliminary
##molecule_type mRNA
##residues       1-714 ##label DUM
CLASSIFICATION   ##superfamily MAM homology; astracin homology
```

```

FEATURE
43-3229      #domain actacin homology #label AST\
232-401      #domain MAM homology #label MAM
SUMMARY
#length 714  #molecular-weight 80703  #checksum 3852

```

Query Match 3.8%; Score 139; DB 11; Length 714;  
Best Local Similarity 32.9%; Pred. No. 3.24e-04;  
Matches 28; Conservative 21; Mismatches 30; Indels 6; Gaps 6;

[illegible]

ENTRY	15
TITLE	A4472 #type complete
ORGANISM	photomorphogenesis repressor COP1 - Arabidopsis thaliana
DATE	#formal_name Arabidopsis thaliana #common_name mouse-ear cress
	30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change

15

ACCESSIONS A44272

A44272

Deng,

Feld

cell (COP1,

both

nces MU

Columb  
A44272

၃

type n 1

ference

```

8
length

```

•

Conse

српска

—

## CSACKNI

Search completed: Tue Dec 10 07:08:25 1996  
Job time : 71 secs.

\*\*\*\*\*  
Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.  
\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPerch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 10 07:06:29 1996; MasPar time 13.19 Seconds

660,452 Million cell updates/sec

Tabular output not generated.

Title: >US-08-446-915-4  
Description: (1-501) from US08446915.pep  
Perfect Score: 3702

Sequence: 1 MAASVTSPGSELLQPGFS.....NSTYRDNAIFKAVIDLTLGL 501

Scoring table: PAM 150  
Gap 11

Searched: 49340 seqs, 17385503 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: swiss-prot32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9

Statistics: Mean 50.182; Variance 93.568; scale 0.536

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
	1	3702	100.0	501	8	TNF2_MOUSE	TNF RECEPTOR ASSOCIAT	0.00e+00
	2	972	26.3	409	8	TNF1_MOUSE	TNF RECEPTOR ASSOCIAT	1.75e-178
	3	229	6.2	458	2	DG17_DICD1	DG17 PROTEIN.	1.30e-21
	4	162	4.4	760	5	MEPA_MOUSE	MEPRIN A ALPHA-SUBUNI	7.47e-10
	5	158	4.3	487	6	RA18_YEAST	DNA REPAIR PROTEIN RA	3.37e-09
	6	157	4.2	668	5	MEPB_RAT	MEPRIN A BETA-SUBUNIT	4.90e-09
	7	153	4.1	1483	2	CYP1_YEAST	CYP1 ACTIVATORY PROTE	2.17e-08
	8	141	3.8	377	7	RINI_HUMAN	RING1 PROTEIN.	1.69e-06
	9	135	3.6	675	2	COP1_ARATH	COP1 REGULATORY PROTE	1.40e-05
	10	126	3.4	647	9	ZG48_XENLA	GASTRULA ZINC FINGER	3.04e-04
	11	120	3.2	160	9	YNN1_CAEEL	HYPOTHEITICAL 18.7 KD	2.21e-03
	12	119	3.2	227	8	TPMS_HUMAN	TROPOMYOSIN ALPHA CHA	3.06e-03
	13	120	3.2	1863	1	BRC1_HUMAN	BREAST CANCER TYPE 1	2.21e-03

14	113	3.1	400	6	OMPA_THEMA	OUTER MEMBRANE PROTEI	2.07e-02
15	114	3.1	501	8	UVS2_NEICR	UVS-2 PROTEIN.	1.51e-02
16	110	3.0	529	8	ZG20_XENLA	GASTRULA ZINC FINGER	5.24e-02
17	111	3.0	1040	7	RAG1_MOUSE	V(D) J RECOMBINATION A	3.85e-02
18	112	3.0	1042	7	RAG1_RABIT	V(D) J RECOMBINATION A	2.82e-02
19	110	3.0	1043	7	RAG1_HUMAN	V(D) J RECOMBINATION A	5.24e-02
20	106	2.9	389	5	MYSP_SCHLA	PARAMYOSIN (FRAGMENT)	1.76e-01
21	109	2.9	393	5	MLZ4_MOUSE	ZINC FINGER PROTEIN M	7.12e-02
22	106	2.9	458	4	IE31_TORCA	TYPE III INTERMEDIATE	1.76e-01
23	107	2.9	879	5	MYSP_DROME	PARAMYOSIN, LONG FORM	1.31e-01
24	104	2.8	248	8	TPM3_CHICK	TROPOMYOSIN BETA 3, F	3.19e-01
25	104	2.8	284	8	TPM1_HUMAN	TROPOMYOSIN ALPHA CHA	3.19e-01
26	104	2.8	284	8	TPM1_CHICK	TROPOMYOSIN 1, SMOOTH	3.19e-01
27	103	2.8	284	8	TPM8_CHICK	TROPOMYOSIN BETA CHAI	4.28e-01
28	102	2.8	284	8	TPM1_CLOIN	TROPOMYOSIN, SMOOTH M	5.73e-01
29	104	2.8	309	5	MYSA_MESNU	MYOSIN HEAVY CHAIN, C	3.19e-01
30	102	2.8	430	1	AROA_STAUD	3-PHOSPHOSHIKIMATE 1-	5.72e-01
31	104	2.8	458	8	VIMI_XENLA	VIMENTINS 1 AND 2.	3.19e-01
32	105	2.8	463	8	VIM4_XENLA	VIMENTIN 4.	2.37e-01
33	104	2.8	624	1	A33_PLENA	ZINC-BINDING PROTEIN	2.37e-01
34	105	2.8	648	2	CH44_YEAST	CH44 ACTIVATORY PROTE	2.37e-01
35	102	2.8	811	4	HSP7_YEAST	MITOCHONDRIAL HEAT SH	5.73e-01
36	102	2.8	1959	5	MYSN_CHICK	MYOSIN HEAVY CHAIN, N	5.73e-01
37	101	2.7	284	8	TPM8_RABIT	TROPOMYOSIN BETA CHAI	7.64e-01
38	101	2.7	284	8	TPM8_HUMAN	TROPOMYOSIN BETA CHAI	7.64e-01
39	100	2.7	284	8	TPM4_XENLA	GASTRULA ZINC FINGER	1.02e+00
40	101	2.7	337	9	ZG26_XENLA	GASTRULA ZINC FINGER	7.64e-01
41	101	2.7	403	4	KICS_MOUSE	KERATIN, TYPE I CYTOS	7.64e-01
42	101	2.7	454	8	TNR1_MOUSE	TUMOR NECROSIS FACTOR	7.64e-01
43	101	2.7	465	8	VIME_HUMAN	VIMENTIN.	7.64e-01
44	101	2.7	1314	1	ADR6_YEAST	TRANSCRIPTION FACTOR	7.64e-01
45	101	2.7	1978	5	MYSG_CHICK	MYOSIN HEAVY CHAIN, G	7.64e-01

#### ALIGNMENTS

RESULT	ID	TNF2_MOUSE	STANDARD;	PRT;	501 AA.
AC	P39429;				
DT	01-FEB-1995 (REL. 31, CREATED)				
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)				
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)				
DE	TNF RECEPTOR ASSOCIATED FACTOR 2 (TNFA2).				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 94349371.				
RA	ROTHE M., WONG S.C., HENZEL W.J., GOEDEDEL D.V.;				
RL	CELL 78:681-692(1994).				
CC	-1- FUNCTION: SIGNAL TRANSDUCER ASSOCIATED WITH THE CYTOPLASMIC DOMAIN				
CC	OF THE 75 KD TUMOR NECROSIS FACTOR RECEPTOR (TNF-R2).				
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.				
CC	-1- SUBUNIT: HETERODIMER OF TNFA1 AND TNFA2.				
CC	-1- SIMILARITY: CONTRAINS A C3HC4-CLASS ZINC FINGER.				
DR	EMBL; L35303; L35303.				
DR	PROSITE; PS00518; ZINC_FINGER_C3HC4.				
KM	ZINC_FINGER; COILED COIL.				
FT	ZN FING 34 72 C3HC4-TYPE.				
SQ	SEQUENCE 501 AA; 56026 MW; 1264825 CN;				
Query Match	100.0%; Score 3702; DB 8; Length 501;				
Best Local Similarity	100.0%; Pred. No. 0.00e+00;				

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	1	maasrtspgslelilpgfsaklllgttrleakyjlsacknlllrpbgaqchryrscflts	60
Qy	1	MAAaSVtSGSLlELlPGfSKtLLGTRlEaKYLCSACKnTLlRPFQaCGhRYrSCfLTS	60
Db	61	ilsagpncacryeqlyeejisilseesaafpdaarveesjpaevpndctwqtlke	120
Qy	61	ILSSGPNaCvYEGlYEEGJSILSSSaFPDNaRvEESjPjAVCPNDcCTWqTlKE	120
Db	121	yeschglcpfllileepackqlvrlsekehhegecprelsedqtracpeshvdehve	180
Qy	121	YESCHGLOPFLlLEePCKAGLmLSKEHHeGECPReLSedQTRaCPeSHVDeHVE	180
Db	181	vcpbfltcdegcqkklipretfqdhvrasacervlcrfhvcgsamevenlgdheqlqr	240
Qy	181	VCPBFLtCdEGCqKklIPrETfQdHvRASaCERVlCRfhVCgsaMeVeNLGDHeQLQR	240
Db	241	rehlallllesflseagaepqtlngvypelldtrcqillegklatfeniwcvlmevexarvta	300
Qy	241	REHLALLSLfLEaQSPeGTLNgVPeLLDTRCqILlEGKlATFeNiWCvLMEVEXArVtA	300
Db	301	eacsrqtrlddqkileasrhvkvgqterisqldkalamadeqkvselevesydygrfiknsd	360
Qy	301	EACSRQHRldDQKlEaSRhVKVgQTeRISqLDKaMaDeQKVSeLeVSYDYGrFIKNSD	360
Db	361	ftckrqeavagrpaisfaftylerisykmclrryngdqtgrghslffvymkpnda	420
Qy	361	FTCKRQEAVaGRPaISfaFtyLERISYKMcLRRyNgDQtGRGHSLffVYMKPNDa	420
Db	421	llqwfngqvltlmlldhnmrehvldafipdvtsssfqrpvsdmiiasgcplfcopvsknea	480
Qy	421	LLQWfNgQVtLmLLDhNMRhVldAFIPdVtSSsFQRpVSdMiIASGCPlFCOpVSKNEA	480
Db	481	kneyvrdaisfikaisvdlfcgl	501
Qy	481	KNeYvRDAISfIKaISVdlfCGL	501

ID	RESULT	2	STANDARD	PRT	409 AA.
AC	P39428;				
DT	01-FEB-1995 (REL. 31, CREATED)				
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)				
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)				
DE	TNF RECEPTOR ASSOCIATED FACTOR 1 (TRAF1).				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
NC	EUTHERIA; RODENTIA.				
LN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 123-135 AND 390-402.				
RX	MEDLINE; 94349371.				
RA	ROTHE M., MONG S.C., HENZEL W.J., GOEDEL D.V.;				
RL	CELL 78:681-692(1994).				
CC	-1- FUNCTION: SIGNAL TRANSDUCER ASSOCIATED WITH THE CYTOPLASMIC DOMAIN				
CC	-1- OF THE 75 KD TUMOR NECROSIS FACTOR RECEPTOR (TNF-R2).				
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.				
CC	-1- SUBUNIT: HETERODIMER OF TRAF1 AND TRAF2.				
DR	EMBL; L35302; L35302.				
KM	COILED COIL.				
SQ	SEQUENCE 409 AA; 45464 MW; 857796 CN;				

Query Match	26.3%;	Score 972;	DB 8;	Length 409;
Best Local Similarity	54.0%;	Pred. No. 1.75e-178;		

Matches 121; Conservative 61; Mismatches 41; Indels 1; Gaps 1;

[illegible]

RESULT	3	STANDARD	PRT	458 AA.
ID	DG17 DICI			
AC	P11467;			
DT	01-OCT-1989 (REL. 12, CREATED)			
DT	01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)			
DT	01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)			
DE	DG17 PROTEIN.			
GN	ZFAA OR DG17.			
OS	DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).			
OC	EUKARYOTA; PROTOZOA; SARCOMASTICOPHORA; SARCODINA; RHIZOPODA;			
OC	EUMLCETOZA; DICTYOSTELIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 88142840.			
RA	DRISCOLL D.M., WILLIAMS J.G.;			
RA	MOL. CELL. BIOL. 7:4482-4489(1987).			
CC	-I- THE EXPRESSION OF DG17 PROTEIN IS DEVELOPMENTALLY REGULATED			
CC	-I- INDUCTION: BY CAMP DURING AGGREGATION.			
DR	EMBL; M18106; M18106.			
DR	PIR; A29361; A29361.			
DR	DICTYDB; DD0210; ZFAA.			
KM	DEVELOPMENTAL PROTEIN; ZINC-FINGER.			
FT	SIMILAR	25	67	TO DR0SOPHILA SINA (AA 71-108).
FT	ZN FING	27	66	POTENTIAL.
FT	ZN FING	178	198	POTENTIAL.
SO	SEQUENCE	458 AA;	53015 MD;	1107348 CN;

Query Match	6.2%;	Score 229;	DB 2;	length 458;
Best Local Similarity	24.8%;	Pred. No. 1.30e-21;		
Matches	64;	Mismatches 103;	Incls 23;	Gaps 22;

[illegible]







```

RA LOVERING R., HANSON I.M., BORDEN K.L.B., MARTIN S., O'REILLY N.J.,  

RA EVAN G.I., RAHMAN D., PAPPIN D.J.C., TROMSDALE J., FREEMONT P.S.;  

RL PROC. NATL. ACAD. SCI. U.S.A. 90:2112-2116(1993).  

CC -I- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  

CC -I- SIMILARITY: CONTAINS A CHC4-CLASS ZINC FINGER.  

DR EMBL; 214000; 214000.  

DR PIR; A47380; A47380.  

DR HSP; p28990; 1CHC.  

DR PROSITE; PS00518; ZINC FINGER C3HC4.  

KM ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.  

FT ZN_FING 19 58 C3HC4-TYPE.  

FT DOMAIN 176 231 GLY-RICH.  

FT DOMAIN 285 348 GLY-RICH.  

FT DOMAIN 172 175 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  

SQ SEQUENCE 377 AA; 39145 MW; 71628 CN;  

Db 13 lhseimpcidmlkntmkcclhfscdcivrlsqnkeptic 58  

| : :: | :: | : : | ||| | : | | |::|  

Qy 28 lEAKYLCSACKRIILRRPQA-QGGRYSCECLTSLSSQGNCAC 72  

RESULT 9  

ID ID COP1_ARATH STANDARD; PRT; 675 AA.  

AC P43254;  

DT 01-NOV-1995 (REL. 32, CREATED)  

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  

DE COP1 REGULATORY PROTEIN (FUSCA PROTEIN FUS1).  

CN COP1 OR FUS1.  

OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  

OC EUKARYOTA; PLANTA; EMPHYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;  

OC CAPRARIACEAE; CRUCIFERAE.  

RN [1]  

RP SEQUENCE FROM N.A.  

RC STRAIN=CV, COLUMBIA; TISSUE=WHOLE SEEDLING;  

RX MEDLINE; 93046683.  

RA DENG X.-W., MATSUI M., WEI N., WAGNER D., CHU A.M., FELDMANN K.A.,  

RA QVALI P.H.;  

RA CELL 71:791-801(1992).  

RL -I- FUNCTION: ACTS AS A REPRESSOR OF PHOTOMORPHOGENESIS IN DARKNESS,  

AND LIGHT STIMULI ABROGATE THIS SUPPRESSIVE ACTION. COULD FUNCTION  

AS A NEGATIVE TRANSCRIPTIONAL REGULATOR CAPABLE OF DIRECT  

INTERACTION WITH COMPONENTS OF THE G PROTEIN SIGNALING PATHWAY.  

CC -I- SUBCELLULAR LOCATION: NUCLEAR.  

CC -I- SIMILARITY: CONTAINS A CHC4-CLASS ZINC FINGER.  

CC -I- SIMILARITY: BELONGS TO THE BETA TRANSDUCCIN FAMILY; CONTAINS  

TRP-ASP DOMAINS.  

DR EMBL; L24437; L24437.  

DR PROSITE; PS00518; ZINC_FINGER_C3HC4.  

DR PROSITE; PS00678; G_BETA_REPEATS.  

KM ZINC-FINGER; NUCLEAR PROTEIN; REPEAT.  

FT ZN_FING 52 89 CHC4-TYPE.  

SQ SEQUENCE 675 AA; 76260 MW; 2370819 CN;  

Query Match 3.6%; Score 141; DB 7; Length 377;  

Best Local Similarity 31.4%; Pred. No. 1,69e-06;  

Matches 15; Conservative 14; Mismatches 16; Indels 1; Gaps 1  


```

```

RA  LOVERING R., HANSON I.M., BORDEN K.L.B., MARTIN S., O'REILLY N.J.,
RA  EVAN G.I., RAHMAN D., PAPPIN D.J.C., TROMSDALE J., FREEMONT P.S.;
RL  PROC. NATL. ACAD. SCI. U.S.A. 90:2112-2116(1993).
CC  -I- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC  -I- SIMILARITY: CONTAINS A CHC4-CLASS ZINC FINGER.
DR  EMBL; 214000; 214000.
DR  PIR; A47380; A47380.
DR  HSP; P28990; 1CHC.
DR  PROSITE; PS00518; ZINC FINGER CHC4.
KM  ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.
FT  ZN FING 19 58 CHC4-TYPE.
FT  DOMAIN 176 231 GLY-RICH.
FT  DOMAIN 285 348 GLY-RICH.
FT  DOMAIN 172 175 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ  SEQUENCE 377 AA; 39145 MW; 716288 CN;

Db  13 lhseimpcidmlkntkctkclrfscdvtalrsgnkeptic 58
Qy  28 lEAKYLCsAKNIILRRPQA-QGCHRYSCFCLTSLSSGQNCAAC 72

RESULT 9
ID  COP1_ARATH STANDARD; PRT; 675 AA.
AC  P43254;
DT  01-NOV-1995 (REL. 32, CREATED)
DT  01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT  01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE  COP1 REGULATORY PROTEIN (FUSCA PROTEIN FUS1).
GN  COP1 OR FUS1.
OS  ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC  EUKARYOTA; PLANTA; EMNOROPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC  CAPRAALES; CRUCIFERAE.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=CV, COLUMBIA; TISSUE=WHOLE SEEDLING;
RX  MEDLINE; 93046683.
RA  DENG X.-W., MATSUI M., WEI N., WAGNER D., CHU A.M., FELDMANN K.A.,
RA  QVALI P.H.;
CELL 71:791-801(1992).
RL  -I- FUNCTION: ACTS AS A REPRESSOR OF PHOTOMORPHOGENESIS IN DARKNESS,
CC  AND LIGHT STIMULI ABROGATE THIS SUPPRESSIVE ACTION. COULD FUNCTION
CC  AS A NEGATIVE TRANSCRIPTIONAL REGULATOR CAPABLE OF DIRECT
CC  INTERACTION WITH COMPONENTS OF THE G PROTEIN SIGNALING PATHWAY.
CC  -I- SUBCELLULAR LOCATION: NUCLEAR.
CC  -I- SIMILARITY: CONTAINS A CHC4-CLASS ZINC FINGER.
CC  -I- SIMILARITY: BELONGS TO THE BETA TRANSDUCIN FAMILY; CONTAINS
CC  TRP-ASP DOMAINS.
DR  EMBL; L24437; L24437.
DR  PROSITE; PS00518; ZINC_FINGER_CHC4.
DR  PROSITE; PS00678; G_BETA_REPEATS.
KM  ZINC-FINGER; NUCLEAR PROTEIN; REPEAT.
FT  ZN FING 52 89 CHC4-TYPE.
SQ  SEQUENCE 675 AA; 76260 MW; 2370819 CN;

Query Match 3.68; Score 141; DB 7; Length 377;
Best Local Similarity 31.48; Pred. No. 1,69e-06;
Matches 15; Conservative 14; Mismatches 16; Indels 1; Gaps 1.

Db  46 ldkallcpicmgikdflfcaqhsfcymcithlnks-dcpoc 89
SQ  SEQUENCE 461 AA; 49145 MW; 1016288 CN;

```







(a) (b) (c) (d) (e) (f)

(7M)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

```

MasParc_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Dec 10 07:08:41 1996; MasPar time 10.68 Seconds
446.743 Million cell updates/sec
Tabular output not generated.

```

```

Title: >US-08-446-915-4
Description: (1-501) from US08446915.pcp
Perfect Score: 3702
Sequence: 1 MAASVTSGSLIETLPQGS.....NSYRDDAIFKALVDTL 501

```

Scoring table: PAM 150  
Gap 11

Searched: 81589 seqs, 9523651 residues

Database: a-geneseq4  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16

Statistics: Mean 35.625; Variance 156.291; scale 0.228

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result	Score	Query Match	Length	ID	Description	Pred. No
1	3700	99.9	501 15	R90578	Mouse TRAF2.	0.00e+00
2	972	26.3	409 15	R90577	Mouse TRAF1.	3.96e-79
3	131	3.5	41 16	R91212	RIN1 zinc finger doma	5.64e-02
4	120	3.2	1863 16	R91208	BRCA1, breast and ova	3.81e-01
5	120	3.2	1863 15	R76641	BRCA1 protein.	3.81e-01
6	111	3.0	41 16	R91210	BRCA1 polypeptide zin	1.74e+00
7	99	2.7	252 1	P94369	Fusion protein congl.	1.24e+01
8	101	2.7	700 15	R89145	Chondroitinase AC.	9.00e+00
9	96	2.6	45 16	R91211	RPT1 zinc finger doma	2.00e+01
10	95	2.6	200 10	R53340	Human/rat chimeric ci	2.34e+01
11	95	2.6	200 10	R53428	Human/rat chimeric ci	2.34e+01
12	95	2.6	200 10	R53429	Human/rat chimeric ci	2.34e+01

13	96	2.6	365	1	P9146	Ret-1 clone cDNA inse	2.40e+01
14	98	2.6	475	3	R15146	Ro/SSA autoantigen.	1.06e+01
15	97	2.6	576	12	R66929	AMM chromosome inv(1	1.71e+01
16	97	2.6	815	12	R66531	AMM chromosome inv(1	1.71e+01
17	97	2.6	885	12	R66530	AMM chromosome inv(1	1.71e+01
18	96	2.6	1093	8	R42818	YMF.	2.00e+01
19	97	2.6	2482	14	R72826	Human mitosh.	1.71e+01
20	92	2.5	263	1	R03348	VP2 sequence for HRV	3.74e+01
21	92	2.5	462	1	R05166	Portion of peptide an	3.74e+01
22	94	2.5	594	6	R34936	CNP-B.	2.74e+01
23	94	2.5	854	12	R66052	Human NMDA receptor s	2.74e+01
24	94	2.5	870	12	R66051	Human NMDA receptor s	2.74e+01
25	94	2.5	908	12	R66053	Human NMDA receptor s	2.74e+01
26	91	2.5	1427	2	R10534	Human 160kD mediator	4.37e+01
27	94	2.5	3084	2	P94158	Sequence of mouse lam	2.74e+01
28	88	2.4	161	7	R37107	N-terminal delta14/C-	6.92e+01
29	88	2.4	187	7	R37105	Delta14 hONTF.	6.92e+01
30	88	2.4	192	6	R32856	Sequence of neurotrop	6.92e+01
31	89	2.4	200	1	R06234	Human sciatic nerve c	5.94e+01
32	88	2.4	200	6	R31950	Sequence of human cll	6.92e+01
33	88	2.4	200	10	R53424	Human/rat chimeric ci	6.92e+01
34	88	2.4	200	11	R58303	Human ciliary neutrop	6.92e+01
35	88	2.4	200	3	R13960	Human CNTF.	6.92e+01
36	88	2.4	200	7	R34431	Sequence of human cll	6.92e+01
37	88	2.4	200	10	R53432	Human/rat chimeric ci	6.92e+01
38	88	2.4	200	7	R37819	Human/rat chimeric ci	6.92e+01
39	88	2.4	200	13	R70146	Sequence of human cll	6.92e+01
40	88	2.4	200	14	R53425	Human/rat chimeric ci	6.92e+01
41	88	2.4	200	14	R83966	Ciliary neurotrophic	6.92e+01
42	88	2.4	613	1	R06552	Human 5-lipoxygenase	6.92e+01
43	88	2.4	775	3	P50121	Sequence of a polypep	5.94e+01
44	89	2.4	775	2	P70417	Polypeptide with IgE	6.92e+01
45	90	2.4	3685	1	P90290	Human Duchenne muscul	5.10e+01

## ALIGNMENTS

	RESULT	1
ID	R90578	standard; Protein; 501 AA.
AC	R90578;	
DT	09-APR-1996	(first entry)
DE	Mouse TRAF2.	
KW	TRAF2; tumour necrosis factor receptor associated factor 2;	
KW	TNF; CD40.	
OS	Mus musculus.	
FA	Key	Location/Qualifiers
FT	Domain	272..501
FT	/label= TRAF_domain	
FT	Region	275..351
FT	/label= leucine_zipper_region	
PN	W09533051-A1.	
PD	07-DEC-1995.	
PF	25-MAY-1995; U06639.	
PR	27-MAY-1994; U5-250858.	
PR	28-OCT-1994; U5-331394.	
PR	22-MAY-1995; U5-446915.	
PA	(GETH.) GENENTECH INC.	
PI	Goeddel DV, Rothle M.	
DR	WPI; 96-049310/05.	
DR	N-PSDB; T1262.	
PT	Tumour necrosis factor (TNF) receptor-associated factors - involved	
PT	in mediation of biological activities of TNF and CD40 ligands	
PS	Claim 8; Page 75-76; 116pp; English.	
CC	Mouse tumour necrosis factor receptor associated factor 2 (TRAF2)	







KM	susceptibility gene; protein replacement therapy; diagnosis;
KW	prognosis.
OS	Homo sapiens.
PN	M09605307-AZ.
PD	22-FEB-1996.
PF	11-AUG-1995; U10203.
PR	12-AUG-1994; US-289221.
PR	02-SEP-1994; US-300266.
PR	16-SEP-1994; US-308104.
PR	29-NOV-1994; US-348824.
PR	24-MAR-1995; US-409305.
PR	07-JUN-1995; US-488011.
PR	07-JUN-1995; US-483554.
PR	07-JUN-1995; US-487002.
PA	(MTR1-) MYRIAD GENETICS INC. P (UTAH) UNIV UTAH RES FOUND.
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI	Futreal AP, Goldgar DG, Harshman KD, Kamb A, Miki Y, P Shattuck-Eidens DM, Skolnick MH, Swenson J, Tavtigian SV,
PI	Weissen RM; P WPfj : 96-139703/14.
DR	NP isolated human cancer predisposing gene, BRCA1 - used to develop PT probe, for diagnosis, prognosis and therapy of cancers, partic. PT breast and ovarian cancers PT Example 8; Page 172; 190pp; English.
CC	A sequence (R91210) near the amino terminus of the human BRCA1 polypeptide (R91208) shows considerable homology to zinc finger domains, including RPT1 (R91211) a protein that appears to be a negative regulator of the interleukin-2 receptor in mouse, RINI (R91212) a DNA-binding protein that includes a RING-finger motif related to the zinc finger, and RPT1 (R91213) a putative transcription factor that is the N-terminal domain of the RET oncogene product. CC oncogene product. CC Sequence 41 AA; SQ
DQ	Query Match                3.0%; Score 111; DB 16; Length 41; Best Local Similarity 34.1%; Pred. No. 1.74e+00; Matches 14; Conservative 9; Mismatches 15; Indels 3; Gaps 2;
Bb	1 cpiclelel-kqvptckdhlficfmklklnqdkpsqcplc 40     :   :   :   :   :   :   :   :   :     :   :   :   :   :   :   :   :   :
Oy	34 CSACKNILRRFPQCGRHCYSFCFLTSLLS--SGRNACMC 72
RESULT	
ID	P94369 standard; protein; 252 AA.
BC	P94369;
AC	P94369;
DE	13-JUN-1990 (first entry)
DT	Fusion protein congt. 41KD protein of T. colubriformis.
KW	Parasitic nematode; 41 kD protein; fusion protein.
OS	Trichostrongylus colubriformis.
PN	M08900163-A.
PD	12-JAN-1989.
PF	06-JUL-1988; AU0239.
PR	07-JUL-1987; AU-002940.
PA	(BIOT-) Biotech Austral Pt, (CSIR) Commonwealth Sci and Ind Res Orgn. P Cobon GS, Anstien RA, O'Donnell TJ, Frenkel MD, Kennedy WPK, Savin KW, P Magland BM; P WPfj : 89-039628/05.
DR	N-PDSB; N91214.
PT	Protein derived from parasitic nematode species - used to provide protective immunity against nematode parasites of man and animals. PS Claim 5; page 40; 57pp; English. The sequence encodes a molecule which is only 4AAs shorter than mature H

	CC	contortus 41 kD protein. The 6 AAs on the N-terminus are derived from CC beta-galactosidase. Recombinant organisms contg. DNA encoding the CC protein can be used to produce polypeptide which is capable of protecting CC guinea pigs against parasitism from T. colubriformis. CC See also P94260, and P94366-P94368.
SQ	Sequence	252 AA;
Df	Query Match	2.7%; Score 99; DB 1; Length 252; Best Local Similarity 23.2%; Pred. No. 1.2e+01; Matches 29; Conservative 35; Mismatches 55; Indels 6; Gaps 5;
Dd	102 vdvdedcaetkireadallheteseevarklamveadtleraeaaenkiwel 161   : ::  ::       :: : : :   : ::   : :   :   :	
Dy	227 VETNIADQHEDLRRLHALLLSSFLERASPTGNYGPDL--IQ-RCLLEGTATF 282	
Dz	162 eeelrvgnmksejwekealdredsyegqirtsarllheatreafaesrhklqxe 220	
Dl	283 ENIVCVIANHEVAVVAEACSRQRHRLDKDKTEALSNNYQLERSIGKLDAWADLERK 341  I II II I	
Dm	221 vdile 225	
Dn	342 VSELE 346	
RESDUJ_8	ID R89145 strand; Protein; 700 AA. AC R89145, DT 2APR-1996 (first entry) DE Chondroitinase AC. KM Chondroitinase AC; chondroitinase B; chondroitin lyase; RM chondroitin sulphate. OS Flavobacterium heparinum. FH Key Location/Qualifiers FT Peptide 1..23 FT Protein 24..700 FT /label= Mat.protein PN MO9601894-A1. PD 25-JAN-1996. PF 07-JUL-1995; U08560. PR 08-JUL-1994; US-272247. PA (IBEX-) IBEX TECHNOLOGIES R & D INC. PI Bennett DC, Fink D, Gu K, Laliberte M, Linhardt R; P1 Tkalec AL, Zimmermann J; DR MPI; 96-097624/10. N-PSDB: T10316. PT Purification of chondroitinase AC and chondroitinase B - using osmotic shock and successive chromatography to fractionate the PT proteins PT Claim 7; Page 30-33; 53pp; English. PC Flavobacterium heparinum chondroitinase AC (R89145) and chondroitinase B (R89146) are capable of degrading chondroitin sulphate polysaccharides. The enzymes, which have calculated mol.wts. of 77,169 and 53,563, respectively, can be obt'd. by CC purification from F. heparinum cells or by expression of the CC encoding genes (see T10316 and T10317) in host cells. They are CC useful as tools in determining the role of chondroitin sulphates CC in modulating cellular events, and can also be used to develop therapeutic reagents. SQ Sequence 700 AA;	
Query Match	2.7%; Score 101; DB 15; Length 700; Best Local Similarity 31.9%; Pred. No. 9.00e+00; Matches 22; Conservative 17; Mismatches 23; Indels 7; Gaps 5;	

	RESULT	9
ID	R91211 standard; Peptide; 45 AA.	
AC	R91211;	
DT	05-JUN-1996 (first entry)	
DE	RPT1 zinc finger domain.	
KW	RPT1, BRCA1, breast cancer; ovary cancer; predisposing gene;	
KW	susceptibility gene; protein replacement therapy; diagnosis;	
KW	prognosis.	

	Query Match	2.6%	Score 96;	DB 16;	length 45;
	Best Local Similarity	38.5%	Pred. No. 2.00e+01;		
	Matches	10;	Conservative	7;	Mismatches 9;
					Indels 0;
					Gaps 0;
Db	1 cpicellkpevsadcmhsfractc	26			
	:   :   :   :   :   :				
	:   :   :   :   :   :				
	34 CSACKNI LRRPFOACGCHRYCFCLT	59			

DT 13-DEC-1994 (first entry)

FT	/note="Leu is substd. by Met"
PN	W09400134-A.
PD	28-APR-1994.
PF	08-OCT-1993; U09649.
PR	09-OCT-1992; US-959284.
PA	(RECE-) REGENERON PHARM INC.
PI	Parayotatos N <sub>2</sub>
PI	WPI; 94-151319/18.

Query Match 2.6%; Score 95; DB 10; Length 200;  
 Best Local Similarity 35.6%; Pred. No. 2.34e+01;  
 Matches 21; Conservative 13; Mismatches 21; Indels 4; Gaps 4;  
 Db 65 sembeaer|qen-1qayr-fhvllarledqgvbfpte-g-dfhgaibhlllqvaaaf 119  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 y 224 SEWETEDTADLQRIAREHALLALSSFTLEAQAASGCTINOVGELLACQCTILEDKTAT 282

RESULT	11
ID	R53428 standard; protein; 200 AA.
AC	R53428;
DE	13-DEC-1994 (first entry)
DT	Human/rat chimeric ciliary neurotrophic factor RBN218.
KW	Human ciliary neurotrophic factor; hCNTF; mutant; variant;
KW	receptor binding; enhance; nervous system disorder; treatment;
KM	chimeric protein; electrophoretic mobility.
OS	Chimeric Homo sapiens.
OS	Chimeric Rattus sp.
PH	Key
PH	Location/Qualifiers
FT	Misc difference 35
FT	/note="Thr is subst. by Met"
FT	Misc difference 33
FT	/note="Ala is subst. by Val"
FT	Misc difference 56
FT	/note="Met is subst. by Val"
FT	Misc difference 63
FT	/note="Gln is subst. by Arg"



Dec 10 06:53

US-08-446-915-4.rag

15

AC R15148;  
DT 14-FEB-1992 (first entry)  
DE Ro/SSA autoantigen.  
KW Autoantibody; autoantigen; SLE; systemic lupus erythematosus.  
OS Homo sapiens.  
PN W0911711-A.  
PD 14-NOV-1991.  
PE 07-MAY-1991; U03139.  
PR 07-MAY-1990; US-520270.  
PT (OKLA-) OKLAHOMA MED RES FO.  
PI Frank MB, Itoh K;  
DR WPI; 91-353712/48.  
DR N-PSDB; Q14798.  
PT DNA encoding an Ro-SSA autoantigen - useful for diagnosing  
PT auto-immune disorders or presence of auto-antibodies  
PS Disclosure; Fig 2; 41pp; English.  
CC A cDNA library (from human thymus mRNA) in lambda gtl1 was screened  
CC with serum from a patient having systemic lupus erythematosus. Two  
CC clones were reactive with sera (from a panel of lupus patients)  
CC which cont'd. autoantibodies against 52 kD protein.  
CC Both the cDNA and the protein expressed from it, or portions of it,  
CC are useful as diagnostic agents in the identification of patients  
CC having autoantibodies and in the identification and analysis of  
CC the structural and functional properties of the autoantigen and for  
CC application in immunotherapeutic regimens.  
SQ Sequence 475 AA;

	Qwey Match	2.68;	Score 98; DB 3;	length 475;
	Best Local Similarity	28.2%;	Pred. No. 1.46e+01;	
Matches	11;	Conservative	10;	Mismatches 18; Indels 0; Gaps 0;
Dn	16 cpctcdptfvepsiecgchscgcacigqykkggswcpcvcc	54		
Oy	: : : : :   : : : : :	:		:
34	CACGAKILRRPFOAGCHHRCFCTLTSLSSGGQCACAAC	72		

Accession	Gene	Protein	Location/Qualifiers
AC	R66929	standard; Protein; 576 AA.	
DE	R66929	01-SEP-1995 (first entry)	
DT	AMML	chromosome inv(16) product.	
KM	AMML	acute myelomonocytic leukemia; chromosome-16; inversion;	
KM	inv(16)	CBF-beta; CBFb gene; transcription factor; myosin; MYH11;	
KM	SMHHC		
OS	Homo sapiens.		
FH	Key		
FT	Peptide	1..164	
FT	/label= CBFb		
FT	Peptide	165..576	
FT	/label= MYH11		
PN	W09504067-A.		
PD	09-FEB-1995.		
PF	26-JUL-1994; U08530.		
PR	29-JUL-1993; U0-099869.		
PA	(UNMI ) UNIV MICHIGAN.		
PA	(TEXA ) UNIV TEXAS SYSTEM.		
PI	Claxton D, Collins FS, Liu P, Siciliano MJ;		
DR	WPI; 95-082178/11.		
DR	N-PSDB; Q84588.		
PT	Novel DNA spanning the pericentric inversion of chromosome 16 -		
PT	for the screening of acute myeloid Leukemia		
CC	Claim 4; Page 28-30; 78pp; English.		
CC	PCR was performed on total cellular RNA from 5 AMML patients having		
CC	a pericentric inversion of chromosome-16, MdeO subtype. Sequencing		

Dec 10 06:53

US-08-446-915-4 J8g

14

CC showed the *inv(16)* fusion to comprise a sequence from the CBFB  
CC gene, encoding a novel transcription factor, and the MYH11 gene,  
CC encoding smooth muscle myosin heavy chain. In 3 patients, nt 1-492  
CC of the CBFB gene were fused to nt 1921 of MYH11 (shown in  
CC 084588, predicted as sequence in R66929). Probes based on *inv(16)*  
CC can be used for diagnosis of AMLt.  
CC Sequence 5/6 AA;  
CC

**Query Match** 2.68; Score 97; DB 12; Length 576;  
Best Local Similarity 29.6%; Pred. No. 1.71e+01;  
**Matches** 16; Conservative 20; Mismatches 15; Indels 3; Gaps 3.

Search completed: Tue Dec 10 07:09:39 1996  
Job time : 58 secs.

Dec 10 06:49

US-08-446-915-2.rng

1

US-08-446-915-2 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPerch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 10 07:04:40 1996; MsePar time 9.40 Seconds

414.306 Million cell updates/sec

Tabular output not generated.

Title: >US-08-446-915-2

Description: (1-409) from US08446915.pep

Perfect Score: 2945

Sequence: 1 MASSSADPENEQFGCPAP.....KNAVYKDDTMEKCIQVTSA 409

Scoring table: PAM 150

Gap 11

Searched: 81589 seqs, 9523651 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-geneseq24

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16

Statistics: Mean 34.957; Variance 160.267; scale 0.218

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description	Pred. No
	1	2945	100.0	409	15	R90577	Mouse TRAF1.	1.47e-259	
	2	972	33.0	501	15	R90578	Mouse TRAF2.	7.66e-76	
	3	112	3.8	870	6	R31348	Jaagsiekte retrovirus	1.77e+00	
	4	113	3.8	1427	2	R10534	Human 160KD mediator	1.51e+00	
	5	105	3.6	383	12	R62656	Petunia Ph6 gene prod	5.43e+00	
	6	103	3.5	192	12	R62757	SetU2 sequence.	7.45e+00	
	7	103	3.5	509	5	R30429	Major Yo paraneoplast	7.45e+00	
	8	99	3.4	77	6	R32698	SSP polypeptide produ	1.39e+01	
	9	99	3.4	77	16	R78258	Part of Major Yo para	1.39e+01	
	10	101	3.4	430	5	R26173	Ethylene response (ET	1.02e+01	
	11	101	3.4	738	15	R69851	Ethylene response (ET	1.02e+01	
	12	101	3.4	738	15	R69852	Ethylene response (ET	1.02e+01	

Dec 10 06:49

US-08-446-915-2.rng

2

13	101	3.4	738	15	R69853	Ethylene response (ET	1.02e+01
14	101	3.4	738	15	R69849	Ethylene response (ET	1.02e+01
15	101	3.4	738	15	R69850	Ethylene response (ET	1.02e+01
16	100	3.4	1093	8	R42818	TNF.	1.19e+01
17	94	3.2	405	11	R57498	Human variant cortico	3.00e+01
18	93	3.2	405	11	R57498	Human corticosteroid	3.49e+01
19	93	3.2	405	11	R61174	Corticosteroid bindin	3.49e+01
20	93	3.2	444	16	R86801	T. thermophilus gamma	2.57e+01
21	95	3.2	1618	5	R27205	Human nestin.	2.57e+01
22	95	3.2	1618	11	R60127	Human nestin protein	2.57e+01
23	92	3.1	235	1	P94260	41kD protein of T. co	4.06e+01
24	92	3.1	320	1	P94366	41kD protein of T. co	4.06e+01
25	92	3.1	362	8	R41536	Preprokallin 2 toxin	4.06e+01
26	90	3.1	394	16	R94559	Human Gs alpha protei	5.47e+01
27	92	3.1	462	1	R05766	Portion of peptide an	4.06e+01
28	91	3.1	476	8	R43563	Hyaluronan receptor.	4.77e+01
29	90	3.1	648	16	R94906	RAP-1 radiation prote	5.47e+01
30	92	3.1	955	11	R57365	K39 polypeptide of Le	4.06e+01
31	90	3.1	1147	8	R41199	CAI antigen.	5.47e+01
32	91	3.1	15281	8	R44929	T. ntvuum Cyclosporin	4.77e+01
33	88	3.0	239	2	R04117	ORF2 of Enod2a genom	7.37e+01
34	88	3.0	239	2	R04118	ORF1 of Enod2b genom	7.37e+01
35	88	3.0	433	2	R08335	CD4.	7.37e+01
36	88	3.0	435	1	P90992	Human CD4 antigen	7.37e+01
37	88	3.0	435	2	P91289	Sequence of a segment	7.37e+01
38	89	3.0	615	15	R74630	Tomato TGEV1 ethylen	6.35e+01
39	88	3.0	816	12	R66931	AMM1 chromosome inv(1	7.37e+01
40	88	3.0	885	12	R66930	AMM1 chromosome inv(1	7.37e+01
41	88	3.0	897	16	R92751	Marine EGF receptor s	7.37e+01
42	88	3.0	1365	6	R36780	KRES.	7.37e+01
43	88	3.0	1764	1	P91672	Primary amino acid se	6.35e+01
44	89	3.0	2101	8	R47173	Sequence of the inner	6.35e+01
45	88	3.0	5035	5	R25450	MH mutant porcine rya	7.37e+01

#### ALIGNMENTS

RESULT	1
ID	R90577 standard; Protein; 409 AA.
AC	R90577;
DT	09-APR-1996 (first entry)
DE	Mouse TRAF1.
KW	TRAF1; tumour necrosis factor receptor associated factor 1;
KW	TNF; CD40.
OS	Mus musculus.
FH	Key
FT	Domain
FT	/label= TRAF_domain
FT	/label= Leucine_zipper_region
PN	MO9533051-AL.
PD	07-DEC-1995.
PF	25-MAY-1995; U06639.
PR	27-MAY-1994; US-250858.
PR	28-OCT-1994; US-331394.
PR	22-MAY-1995; US-446915.
PA	(GETH) GENENTECH INC.
PI	Goeddel DV; Roche M
DR	WPI; 96-049310/05.
DR	N-PSDB; T1261.
PT	Tumour necrosis factor (TNF) receptor-associated factors - involved
PT	in mediation of biological activities of TNF and CD40 ligands
PS	Claim 8; Page 71-72; 116pp; English.
CC	Mouse tumour necrosis factor receptor associated factor 1 (TRAF1)



CC See also R31346-7, R31349 and Q35153-Q35155.  
SQ Sequence 870 AA;

Query Match 3.8%; Score 112; DB 6; Length 870;  
Best Local Similarity 24.7%; Pred. No. 1.77e+00;

Matches 24; Conservative 27; Mismatches 42; Indels 4; Gaps 4;

Db 187 qlylvhnddllahtcdhl-lygaf-eilkqlhslngtviadekigtffpyngfsfjy 244  
:| | : | | | : | | : | | : | | : | | : | | : | | : | |

Qy 168 ELALQHLVKEKLLAQLEEKLNPFNVAVLNKEVEASHLALASTHQSQLDREHL-LSTLE 226

Db 245 pwyntqlvklqfcdh-lktlndfgkljgdmwlrpyl 280  
|| : | | | : | | : | | : | | : | |

Qy 227 QRVEIQAQTIAQKQDVQKGLKLSRLMEASFDCTFL 263

## RESULT 4

ID R10534 standard; Protein; 1427 AA.

AC R10534;

DT 12-APR-1991 (first entry)

DE Human 160kD mediator of inflammation protein.

KM Mediator of inflammation; cytokine; Hodgkin's lymphoma; MRP-160.

OS Homo sapiens.

PN EP-412050-A.

PD 06-FEB-1991.

PF 26-JUN-1990; 810481.

PR 05-JUL-1989; GB-015414.

PA (CIBA ) CIBA GEIGY AG.

PI Odink KG, Tarcsay L, Bruggen J, Wiesendanger W, Cerletti N;

PI Sorg C, Dewolf-Peters C, Delabie J;

DR WPI; 91-038913/06.

DR N-PSDB; Q10378.

PT 160 kD human polypeptide mediator or precursor of inflammation -

PT polyclonal or monoclonal antibodies to polypeptide treat and

PT diagnose chronic inflammation and hodgkins lymphoma

PS Claim 3; Page 32; 47pp; English.

CC The protein is a cytokine used to treat chronic inflammatory

CC conditions. It is prepared by chromatographically purifying an

CC optionally pre-purified cell extract, cell supernatant or cell

CC filtrate of stimulated normal human leucocytes or human embryonic

CC epithelial lung cells. Alternatively, the protein can be produced

CC by microorganisms or continuous mammalian cell lines, transformed

CC with plasmids encoding MRP-160. The invention also covers the

CC polypeptide fragment from amino acids 878-1427 and derivatives of

CC the protein in which the amino and/or hydroxyl functions are

CC glycosylated or acylated and have mol. wt.s of 190 and 140kD,

CC respectively.

SQ Sequence 1427 AA;

Query Match 3.8%; Score 113; DB 2; Length 1427;  
Best Local Similarity 25.0%; Pred. No. 1.51e+00;

Matches 24; Conservative 31; Mismatches 34; Indels 7; Gaps 7;

Db 451 tkgdletqtklehari-kelegsl-lfektacklqreledtravtveksrimelekl 508  
:| | : | | : | | : | | : | | : | | : | | : | | : | |

Qy 165 SQEELALQ-HLVKEKLLAQLEEKLNPFNVAVLNKEVEASHLA-LAASIHQSQLDREH 221

Db 509 alrvqe-vaeljrrl-esnkpgdvmslsgis 542  
| : | | : | | : | | : | | : | | : | | : | |

Qy 222 LUSTEQRVEIQAQTIAQKQDVQKGLKLSRLMEAS 257

## RESULT 5

ID R62656 standard; Protein; 383 AA.

AC R62656;  
DT 08-JUN-1995 (first entry)

DE Petunia Ph6 gene product.

KM Vacuolar pH; pH gene; Petunia.

OS Petunia hybrida strain V26.

FH Key Location/Qualifiers

FT Region 192..198

FT /label= Helix-loop-helix structural motif

FT /note= "see also 239..245"

PN W09423561-A.

PD 27-OCT-1994.

PF 15-APR-1994; 004173.

PR 16-APR-1993; US-049282.

PA (DNAP ) DNA PLANT TECHNOLOGY CORP.

PI Chuck GS, Courtney-guterson N, Dooner HK, Keller J;

PI Nijjar CS, Ralston EJ;

DR WPI; 94-341349/42.

DR N-PSDB; Q73000.

PT Petunia Ph gene and constructs containing it - for alteration of

PT vacuolar pH used in the formation of blue flowers

PS Disclosure; Page 42-44; 62pp; English.

CC The V26 strain of Petunia was used. Poly A RNA was isolated from

CC total RNA from flower buds and used to generate a cDNA library in

CC the vector lambda ZapII (Stratagene). The SstI to BamHI fragment at

CC the left hand side of Ac was used to isolate pPet14-1 (contg. pH6

CC cDNA), the sequence for which is claimed (see Q73000 FT). The AA

CC sequence comprises a helix-loop-helix structural motif starting with

CC the sequence NHVLAER (starting at residue 192) and extending to the

CC motif include the myc family of oncogenes, regulators of neuron and

CC muscle development, and regulators of segmentation and organ

CC patterning in Drosophila. In plants, the motif is found in R(S),

CC a protein involved with regulating anthocyanin synthesis in Matze.

CC Comparison of the pH6 and R(S) sequences over the 54 AA helix-loop-

CC helix region detected 57% AA identity. Outside this region only 19%

CC identity was found.

CC Identity was found.

SQ Sequence 383 AA;

Query Match 3.8%; Score 105; DB 12; Length 383;  
Best Local Similarity 27.1%; Pred. No. 5.43e+00;

Matches 23; Conservative 25; Mismatches 32; Indels 5; Gaps 5;

Db 178 frkgsitqpepsqnhvlaerrrreklnerfllrslvptvk-mdkasi-ldtleyvk 235  
:| | : | | : | | : | | : | | : | | : | | : | | : | |

Qy 158 YRAPCCEQGEELALQHLVKEKLLAQ-LEEKLNPFNVAVLNKEVEASHLALASTH-QS 215

Db 236 ql-rkkvgdlearangteatqltkd 259  
|| | : | | : | | : | | : | | : | |

Qy 216 QLDREHLSTLEQRVEIQAQTIAQKD 240

RESULT 6  
ID R62757 standard; Protein; 192 AA.

AC R62757;

DT 26-JUN-1995 (first entry)

DE SetU2 sequence.

KM Salmonella; SetU2; vaccine.

OS Salmonella.

PN W09425598-A.

PD 10-NOV-1994.

PF 26-APR-1994; 1B0207.

PR 26-APR-1993; US-054452.

PA (KING/) KING J.

PA (UYVI-) UNIV VICTORIA INNOVATION & DEV CORP.



Query Match	3.5%	Score 103	DB 12	Length 192
Beet Local Similarity	20.3%	Pred. No. 7.45e+00		
Matches	19	Conservative	29	Mismatches 39
				Indels 4
				Gaps 4
Db	20	ekqngdlassereqemsaetvrlkrqdaedeelllmgldlrltelp-emkyvfiexq	78	
	1	1	1	1
Qy	211	SIHQSLQR-EHLIISTEQRWELADOTIAKQDOVILKELSHIRIMEAEASDTEFLMKRTIV	269	
Db	79	aefsiakalc-rvldqvarnswyrrqgfrilvc	108	
	1	1	1	1
Qy	270	TKRCHESVQGRVTSLESPAFYIAKQTKL-C	299	
	1	1	1	1

	RESULT	7
ID	R30429 standard; Protein; 509 AA.	
AC	R30429;	
DT	27-JAN-1993 (first entry)	
DE	Major Yo paraneoplastic antigen (CDR62) encoded by clones pY1 pY2	
KM	paraneoplastic cerebellar degeneration; PCD; neurological symptoms;	
KW	tumour cancer; brain; therapeutic agent; imaging.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	Region	
FT	Domain	171..192
FT	/note= "leucine zipper"	
FT	Domain	122..170
FT	/note= "super" leucine zipper"	
FT	Misc difference 201	
FT	/note= "SPKK site"	
FT	Domain	205..231
FT	/note= "putative zinc finger"	
FT	Domain	52..80
FT	/note= "highly acidic (pI=3.4) activating domain"	
PN	W09213093-A.	
PD	06-AUG-1992.	
PF	24-JAN-1992; U00620.	
PR	25-JAN-1991; US-646292.	
PA	(SLOK ) SLOAN KETTERING INST CANCER.	
PI	Furneaux HM, Posner JB;	
DR	MPJ: 92-284676/34.	
N-PSDB;	Q27387.	
PT	Major Yo para-neoplastic antigenic polypeptide and its nucleic	
PT	acid - for treatment and diagnosis of para-neoplastic cerebellar	
PT	degeneration	
PC	Claim 5; Fig 7; 5'bp; English.	
CS	This sequence represents the major Yo paraneoplastic	
CC	antigen polypeptide, and was deduced from the cDNA sequences of	
CC	operational clones pY2 (Q27387) and pY1. Anti-Yo antibodies may be	
CC	detected using labelled Yo to see if a patient has paraneoplastic	
CC	cerebellar degeneration (PCD). Labelled antibodies to Yo can be used	
CC	to determine whether a patient is carrying a tumour expressing Yo.	
CC	When coupled to a therapeutic agent, these Abs can be used to	

Query Match 3.5%; Score 103; DB 5; Length 509;  
Best Local Similarity 26.1%; Pred. No. 7,456+00;  
Matches 30; Conservative 30; Mismatches 44; Indels 11; Gaps 10.

RESULT 8  
 ID R32698 standard; Protein; 77 AA.  
 AC R32698;  
 DT 22-JUN-1993 (first entry)  
 DE SSP polypeptide produced from clone 2-9.  
 KW Heptad; plants; custom tailored storage proteins; in vivo;  
 KW expression.  
 OS Synthetic.  
 PN MO9303160-A.  
 PD 18-FEB-1993.  
 PF 07-AUG-1992; 006412.  
 PR 09-AUG-1991; US-743006.  
 PA (DUPRO ) DU PONT DE NEMOURS & CO E I.  
 PI Falco SC, Keeler SJ, Rice JA;  
 DR WPI, 93-076517/09.  
 DR N-PSDB; 037279.  
 PT Synthetic polypeptide(s) contg. specified heptad units -  
 PT expressed in vivo in plants to serve as custom-tailored storage  
 PT proteins with specified aminoacid content  
 PS Disclosure; Page 130; 176pp; English.  
 CC The sequence represents a synthetic heptad polypeptide which can be  
 CC expressed in vivo in plants to serve as a synthetic seed storage  
 CC protein which can be custom-tailored for specific end-user requirements  
 CC The DNA encoding the heptad may be used to transform plants to increase  
 CC the content of partic. amino acids such as lysine or methionine in seeds  
 CC or leaves. See also R31979-86 and R32659-705.  
 SQ Sequence 77 AA;





DE Ethylene response (ETR) gene product.  
KM Ethylene response; recombinant; mutation; decrease; fruit ripening;  
KM abscission; floral senescence.  
OS Arabidopsis thaliana.  
PN W09501439-A2.  
PD 12-JAN-1995.  
PF 30-JUN-1994; U07418.  
PR 01-JUL-1993; US-086555.  
PR 28-JUN-1994; US-263480.  
PA (CALY ) CALIFORNIA INST OF TECHN.  
P1 Bleeker AB, Chang C, Meyerowitz EM;  
DR WPI; 95-061003/08.  
DR N-PSDB; Q85557.  
PT Modified ethylene response (ETR) nucleic acid - useful for producing  
PT transformed plants with decrease in response to ethylene  
PS Claim 1; Page 59-62; 212pp; English.  
CC Q85557 represents the cDNA sequence of the Arabidopsis thaliana  
CC ethylene response (ETR) gene which codes for the ethylene response  
CC protein (R69849). The ETR cDNA or the full gene sequence (Q85556)  
CC can be mutated and linked to a promoter and used to transform plant  
CC cells of fruit bearing plants (in partic. the fruit is a tomato)  
CC resulting in a decrease in response to ethylene. This enables  
CC controlled fruit ripening and delayed floral senescence and  
CC abscission during growth, transport or storage.  
SQ Sequence 738 AA;

Query Match 3.4%; Score 101; DB 15; Length 738;  
Best Local Similarity 26.8%; Pred. No. 1.02e+01;  
Matches 30; Conservative 36; Mismatches 37; Indels 9; Gaps 7;

Db 321 dlmeqvaldlarreaetarndflavmhemtphmhalalsllgeteltpgrl 380  
:|::| | | : | :| ::||| | : : | :||:::| | : |  
Qy 168 ELALQHLVKEKLLAQLEEK-LRFTANIVAVINKEVEAS-H--LALAASIHQSQDREHLL 223  
:  
Db 381 mvetllksnllatlmndvldarledgsqql-elgtfnlhtlfrevnlilk 431  
:  
Qy 224 SLEQ--RVEELQDTLAQKQVLEKLEH-SLRIMEASFD-CTFLMKITVTK 271

RESULT 15  
ID R69850 standard; Protein; 738 AA.  
AC R69850;  
DT 20-MAR-1996 (first entry)  
DE Ethylene response (ETR) mutant protein etrl-1.  
KM Ethylene response; recombinant; mutation; decrease; fruit ripening;  
KM abscission; floral senescence.  
OS Arabidopsis thaliana.  
PN W09501439-A2.  
PD 12-JAN-1995.  
PF 30-JUN-1994; U07418.  
PR 01-JUL-1993; US-086555.  
PR 28-JUN-1994; US-263480.  
PA (CALY ) CALIFORNIA INST OF TECHN.  
P1 Bleeker AB, Chang C, Meyerowitz EM;  
DR WPI; 95-061003/08.  
DR N-PSDB; Q85558.  
PT Modified ethylene response (ETR) nucleic acid - useful for producing  
PT transformed plants with decrease in response to ethylene  
PS Claim 2; Page 65-68; 212pp; English.  
CC R69850-R69853 represent the amino acid sequences of the Arabidopsis  
CC thaliana mutated ethylene response (ETR) proteins etrl-1, etrl-2,  
CC etrl-3 and etrl-4 respectively. The cDNA sequences which encode  
CC these proteins can be linked to a promoter and used to transform  
CC plant cells of fruit bearing plants (in partic. the fruit is a

CC tomato) resulting in a decrease in response to ethylene. This enables  
CC controlled fruit ripening and delayed floral senescence and  
CC abscission during growth, transport or storage.  
SQ Sequence 738 AA;

Query Match 3.4%; Score 101; DB 15; Length 738;  
Best Local Similarity 26.8%; Pred. No. 1.02e+01;  
Matches 30; Conservative 36; Mismatches 37; Indels 9; Gaps 7;

Db 321 dlmeqvaldlarreaetarndflavmhemtphmhalalsllgeteltpgrl 380  
:|::| | | : | :| ::||| | : : | :||:::| | : |  
Qy 168 ELALQHLVKEKLLAQLEEK-LRFTANIVAVINKEVEAS-H--LALAASIHQSQDREHLL 223  
:  
Db 381 mvetllksnllatlmndvldarledgsqql-elgtfnlhtlfrevnlilk 431  
:  
Qy 224 SLEQ--RVEELQDTLAQKQVLEKLEH-SLRIMEASFD-CTFLMKITVTK 271

Search completed: Tue Dec 10 07:06:11 1996  
Job time : 91 secs.